matics

JUL 20 1935
C-2>

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

G. C. EVANS
UNIVERSITY OF CALIFORNIA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
HARRY BATEMAN
W. A. MANNING
MARSTON MORSE
J. R. KLINE
E. P. LANE
HARRY LEVY
ALONZO CHURCH
L. R. FORD
OSCAR ZARISKI

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LVII, Number 3

JULY, 1935

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

The American Journal of Mathematics will appear four times yearly.

The subscription price of the Journal for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the Journal remain on sale.

Papers intended for publication in the Journal may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. Cohen at The Johns Hopkins University.

Subscriptions to the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore, Maryland, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

CYCLOTOMY, HIGHER CONGRUENCES, AND WARING'S PROBLEM II.[1]

By L. E. DICKSON.

PART 2. THE WARING PROBLEM FOR POLYNOMIAL SUMMANDS.

29. *Introduction and summary.* In 1770, E. Waring conjectured that every positive integer N is a sum of 9 integral cubes $\geqq 0$, also that N is a sum of 19 fourth powers, etc. Hardy and Littlewood proved that every sufficiently large N is a sum of

(160) $$s = (\tfrac{1}{2}k - 1)2^{k-1} + k + 5 + \zeta_k$$

integral k-th powers $\geqq 0$, where ζ_k is the greatest integer $\leqq$ the quotient of $(k-2)\log 2 - \log k + \log(k-2)$ by $\log k - \log(k-1)$. Except for very small values of k, ζ_k is quite small compared to (160); for example, $\zeta_{10} = 50$, $\zeta_{28} = 493$.

Waring conjectured also that every N is a sum of a limited number of values of a polynomial in x of degree k. In precise form, this was proved by E. Kampke.[2] But neither writer gave any information as to the number of values needed. For $k = 3$, 9 values suffice.[3] For $k \geqq 4$ the analytic part of the proof that s values of a polynomial suffice for a large N has been made by Miss Humphreys.[4] We here treat the second part of the proof, viz., that *if A is any integer and p is any prime not dividing*[5] *k, then A is congruent modulo p toa sum of n values of the polynomial, where $n < s$ in* (160). We find that

k	3	4	5	6	7	8	9	10
n	4	6	12	24	48	72	144	216
s	9	19	41	87	192	425	949	2113.

If k is one of the even numbers $6, 8, \cdots, 18$, then

$$n = n(k) \leqq 8 \cdot 3^{\frac{1}{2}k-2},$$

which is less than the first term of (160) since $3 < 4$.

[1] Part I of this paper appeared in the current volume of this JOURNAL, pp. 391-424.

[2] *Mathematische Annalen*, Bd. 83 (1921), pp. 85-112.

[3] Dickson, *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 1-12, 739-741; R. D. James, *American Journal of Mathematics*, vol. 56 (1934), pp. 303-315.

[4] *Duke Mathematical Journal*, vol. 1 (1935).

[5] For the case when p divides k, see § 45.

For any odd k, $n(k) \leqq 2n(k-1)$, whence $n(k)$ is much less than the first term of (160) when $n = 7, 9, \cdots, 19$. If we write $s = S + \zeta_k$, then

k	20	22	24	26
n	384,912	1,154,736	57,736,800	173,210,400
S	4,718,617	20,971,547	92,274,717	402,653,215.

Hence $n < s$ for $k < 28$. For $n \geqq 28$, s exceeds a billion, and the actual value of s is of slight interest beyond the fact that there exists an s (Kampke).

30. *Normal polynomials.* We may exclude polynomials $g(x)$ whose values for integers x are all multiples of p, since integers not multiples of p are not represented as a sum of values of $g(x)$. The true Waring problem relates to summands g/p and not to summands g.

By the degree of $g(x)$ modulo p we mean the exponent of the highest power of x whose coefficient c is prime to p. We seek n such that

$$A \equiv \sum_{i=1}^{n} g(x_i) \pmod p \tag{161}$$

has integral solutions x_i for every integer A. We desire that the same n shall serve not only for every A, but for every polynomial $g(x)$ of given degree k modulo p. We write $n = n(k) = n(k, p)$.

Determine d by $cd \equiv 1 \pmod p$. Then dA ranges with A over a complete set of residues modulo p. Hence the problem for (161) reduces to that for $dg(x)$, whose c is $\equiv 1 \pmod p$. If k is prime to p, we take $x = X + z$ and choose z so that the coefficient of X^{k-1} is divisible by p. If C is the constant term of $g(x)$, write $g = H + C$. Then (161) is equivalent to $A - nC \equiv \Sigma H$.

THEOREM 13. *When k is not divisible by p, the problem for* (161) *reduces to the like problem for a* NORMAL *polynomial whose leading coefficient is unity, while the coefficient of x^{k-1} and the constant term are both zero.*

LEMMA 1. *If neither r nor s is divisible by $p > 2$ and if A is any integer, there exist solutions of*

$$rx^2 + sy^2 \equiv A \pmod p.$$

This special case of Theorem 6 is evident since each of rx^2 and $A - sy^2$ takes $1 + \frac{1}{2}(p-1)$ incongruent values and hence have a common value.

When $k = 2$ the only normal polynomial is x^2. Then Lemma 1 with $r = s = 1$ gives

$$n(2, p) = 2, \qquad p > 2. \tag{162}$$

31. *Odd k.* Let k be not divisible by p. By choice of z, $g(x+z)$ becomes

$$f(x)=\sum_{i=0}^{k} c_i x^{k-i}, \qquad c_1 \text{ not divisible by } p.$$
$$f(x)+f(-x)=2c_1x^{k-1}+2c_3x^{k-3}+\cdots=H(x).$$

Take $p>2$. Then $2c_1$ is prime to p. By definition, any A is congruent to a sum of $n(k-1,p)$ values of $H(x)$.

THEOREM 14. *If k is odd and not divisible by $p>2$, then*

$$n(k,p)\leqq 2n(k-1,p).$$

COROLLARY. $n(3,p)\leqq 4$ *if* $p>3$.

32. *Case $k=4$.* We employ the fact that the sum of the fourth powers of $a+d$, $a-d$ and $-2a$ is $2t^2$, the sum of their squares is $2t$, and their sum is zero, where $t=3a^2+d^2$. For $p>2$ every normal polynomial of degree 4 is of the form $f(x)=x^4+2ux^2+vx$. Hence

$$2t^2+2t\cdot 2u=f(a+d)+f(a-d)+f(-2a).$$

The left member is $2(y^2-u^2)$, where $y=t+u$. Employ also a second such identity and add the two. Hence $N=2(y^2+z^2-2u^2)$ is a sum of six values of $f(x)$.

Take $p>3$ and apply Lemma 1. Thus integers y and z may be chosen so that N takes any assigned value modulo p. For the t determined by y, $3a^2+d^2\equiv t \pmod p$ is solvable. Similarly for τ determined by z, $3\alpha^2+\delta^2\equiv\tau \pmod p$ is solvable. This proves

$$n(4,p)\leqq 6 \quad\text{if}\quad p>3. \tag{163}$$

Hence by Theorem 14,

$$n(5,p)\leqq 12 \quad\text{if}\quad p>5. \tag{164}$$

33. LEMMA 2. *For every integer A, there exist*[6] *integral solutions of*

$$\sum_{i=1}^{r} h_i^k\equiv A \pmod p, \tag{165}$$

where r is the g. c. d. of k and $p-1$.

Taking $A=-1$, $k_{r+1}=1$, we obtain

[6] Landau, *Vorlesungen über Zahlentheorie*, I (1927), p. 290. The proof is quite elementary.

LEMMA 3. *For $K = r + 1$, there exist solutions of*

$$(166) \qquad \sum_{i=1}^{K} h_i^k \equiv 0, \text{ not every } h_i \equiv 0 \pmod{p}.$$

34. *Even k.* Let k be not divisible by $p > 2$. As in § 30, it suffices to consider a polynomial of the form $f(x) = x^k + x^{k-1} + \cdots$. By (166),

$$(167) \qquad P(y) = \sum_{i=1}^{K} f(h_i y) \equiv C y^{k-1} + \cdots \pmod{p}, \quad C = \sum_{i=1}^{K} h_i^{k-1}.$$

If C is not divisible by p, we have the result desired. Next, let $C \equiv 0$. By (166) a certain h_j is prime to p. Let $Q(y)$ be derived from $P(y)$ by changing the sign of h_j. If also the leading coefficient of $Q(y)$ is divisible by p, evidently $2h_j^{k-1} \equiv 0$, $k_j \equiv 0 \pmod{p}$, contrary to hypothesis.

THEOREM 15. *Let k be even and not divisible by the odd prime p. Choose K ($\leqq r + 1$) so that* (166) *is solvable. Then*

$$n(k, p) \leqq Kn(k - 1, p).$$

35. *Lemmas, chiefly on congruences.*

LEMMA 4. *If q is prime to $p - 1$, every integer is congruent to a q-th power modulo p.*

Since there are integral solutions of $v(p - 1) + 1 = uq$,

$$x \equiv x(x^{p-1})^v \equiv (x^u)^q \pmod{p}.$$

LEMMA 5. *Let r be the g. c. d. of k and $p - 1$. If each a_i is prime to p,*

$$(168) \qquad a_1 x_1^k + \cdots + a_s x_s^k \equiv c \pmod{p}$$

has the same number of solutions as

$$(169) \qquad a_1 y_1^r + \cdots + a_s y_s^r \equiv c \pmod{p}.$$

Consider any solution of $a_1 z_1 + \cdots + a_s z_s \equiv c \pmod{p}$. For $(i = 1, \cdots, s)$, we shall prove that $x_i^k \equiv z_i$ and $y_i^r \equiv z_i \pmod{p}$ have the same number of roots. This is evident unless z_i is prime to p. Then

$$k \operatorname{Ind} x_i \equiv \operatorname{Ind} z_i \pmod{p - 1}$$

has no root or r roots x_i according as $\operatorname{Ind} z_i$ is not or is divisible by r. The same is true for $r \operatorname{Ind} y_i \equiv \operatorname{Ind} z_i$.

From Theorem 6 and Lemma 5 we obtain

LEMMA 6. *If $p \equiv -1 \pmod{4}$ and if q is prime to $p - 1$, there are exactly $p + 1$ solutions of*

$$x^k + y^k \equiv -1 \pmod p, \qquad k = 2^m q, \quad m \geqq 1.$$

The conditions are satisfied if $p \equiv -1 \pmod{12}$, $q = 3^n$.

LEMMA 7. *If* $p \equiv 1 \pmod 4$ *and if* q *is prime to* $\frac{1}{2}(p-1)$, *there are exactly* $p-1$ *solutions of*

$$x^{2q} + y^{2q} \equiv -1 \pmod p.$$

LEMMA 8. *If* $p > 2$, *there are at most* km *simultaneous*[7] *solutions of*

$$h^k + H^k \equiv -1, \quad h^m + H^m \equiv -1 \pmod p, \qquad m \neq k. \tag{170}$$

Let d be the g. c. d. of k and m. Comparing $(-H^k)^{m/d}$ and $(-H^m)^{k/d}$, we get

$$(h^k + 1)^{m/d} \pm (h^m + 1)^{k/d} \equiv 0, \tag{171}$$

which is not identically $\equiv 0$. It has at most km/d roots. Since d is a linear combination of k, m, we see from (170) that H^d is congruent to a polynomial in h.

LEMMA 9. $n(k, p^e) \leqq p^e - 1$.

We exclude polynomials $f(x)$ all of whose values are multiples of p. As in § 30, we may assume that the constant term of $f(x)$ is zero. Let v denote a value prime to p of $f(x)$.

I. If A is any integer prime to p, $tv \equiv A \pmod{p^e}$ has a solution t, $1 \leqq t < p^e$. Thus A is congruent to a sum of t (equal) values of $f(x)$.

II. If $A \equiv 0 \pmod{p^e}$, then $A \equiv f(0) \pmod{p^e}$.

III. Let $A = p^m a$, $1 \leqq m < e$, a prime to p. By I,

$$a = S + zp^{e-m}, \qquad S = \text{sum of } p^{e-m} - 1 \text{ values of } f(x).$$

Multiply by p^m. Hence A is congruent modulo p^e to

$$p^m S = \text{sum of } p^m(p^{e-m} - 1) < p^e - 1 \text{ values of } f(x).$$

36. *Case* $k = 6$. We shall prove that $n(6, p) \leqq 24$ if $p > 3$.

I. $p \equiv 1 \pmod 4$. Then $h^2 \equiv -1 \pmod p$ is solvable. Hence $h^6 + 1 \equiv 0$. Thus $K = 2$ in (166) and $n(6, p) \leqq 2 \cdot 12$ by Theorem 15 and (164) if $p > 5$. For $p = 5$, use Lemma 9.

II. $p \equiv -1 \pmod{12}$. By Lemma 6, there are exactly $p + 1$ solutions of

[7] If m/d and k/d are both odd, there are at most $2\delta + km - 3dm$ simultaneous solutions, where δ is the number of roots of $z^d \equiv -1 \pmod p$.

$$(172)\qquad h_1^6+h_2^6+h_3^6\equiv 0 \pmod p),\qquad h_3=1.$$

If $p+1>24$, Lemma 8 shows that (172) has a solution for which $h_1^4+h_2^4\not\equiv -1$, and one for which $h_1^2+h_2^2\not\equiv -1 \pmod p$. But if $p=11$ or 23, $n(6,p)\leqq 22$ by Lemma 9.

Consider a normal polynomial

$$(173)\qquad f(x)=x^6+c_4x^4+\cdots+c_1x.$$

Then

$$(174)\qquad P(y)=\sum_{i=1}^{3} f(h_iy)\equiv\sum_{j=1}^{4} C_jM_jy^j \pmod p),\qquad M_j=\sum_{i=1}^{3} h_i^j.$$

If $P(y)$ is not identically $\equiv 0 \pmod p$, $n(6,p)\leqq 3n(4,p)\leqq 18$. Next, let $P(y)$ be identically $\equiv 0$ for all solutions of (172). Since (172) holds also when $h_3=-1$, there are solutions with $M_3\not\equiv 0$, whence $c_3\equiv 0$. Similarly, $M_1\not\equiv 0$, $c_1\equiv 0$. We saw that there are solutions with $M_2\not\equiv 0$, $M_4\not\equiv 0$, whence $c_2\equiv 0$, $c_4\equiv 0$. Hence $f(x)\equiv x^6$ and $n(6,p)=2$ by Lemma 2 or Lemmas 1 and 4.

LEMMA 10. *Except for $p=7, 31, 67, 79, 139, 223$, there exist solutions of $h^6+H^6\equiv -1$ (mod p), if $p\equiv 7$ (mod 12).*

Since $p=6f+1$, f is odd and the number N of solutions is $36(0,3)$ by Theorem 5. By § 19,

$$N=p+1+16A,\quad N=p+1+10A\pm 12B,\quad p=A^2+3B^2,$$

according as 2 is or is not a cubic residue of p. The sign of A was there chosen so that $A\equiv 4 \pmod 6$. By Theorem 7, B is a multiple $3y$ of 3 in the first case; but in the second case, B is prime to 3 and we may choose the sign so that $\pm B\equiv A \pmod 3$.

Let $N=0$. Eliminate p. In the first case,

$$7=\{\tfrac{1}{3}(A+8)\}^2+3y^2=4+3,\ A=-2 \text{ or } -14,\ B^2=9,\ p=31 \text{ or } 223.$$

In the second case, p and 37 are the products of $A\pm B\sqrt{-3}$ and $5-2\sqrt{-3}$ by their conjugates, whence, by multiplication,

$$37p=X^2+3Y^2,\quad X=5A\pm 6B\equiv 2 \pmod 6),\quad Y=-2A\pm 5B,$$

and $Y=3w$, $X+37=3v$, v odd. If $N=0$, $p=-1-2X$,

$$148=v^2+3w^2=121+3\cdot 9 \text{ or } 1+3\cdot 49,$$

whence $p=7$, 139 or 67, 79.

III. $p \equiv 7 \pmod{12}$. First, exclude the six p's in Lemma 10. Then (172) is solvable. There exists an integer e belonging to the exponent 6 modulo p. Hence (172) holds also when h_1 is replaced by eh_1. Hence (172) has solutions for which $M_j \not\equiv 0$, $c_j \equiv 0$ $(j = 1, \cdots, 4$ in turn) in (174).

We have the following solutions of (166) with $K = 4$, $k = 6$.

$$
\begin{array}{llllll}
p = 67, & 1 + 1 + 1 + 2^6 \equiv 0; & & & \\
p = 79, & 1 + 1 + 10 + 67 \equiv 0, & 10 \equiv g^6, & 67 \equiv g^{54}, & g = 29; \\
p = 139, & 1 + 1 + 6 + 131 \equiv 0, & 6 \equiv g^{30}, & 131 \equiv g^{12}, & g = 92; \\
p = 223, & 1 + 4 + 8 + 210 \equiv 0, & 2 \equiv 10^{18}, & 210 \equiv 10^{48}. &
\end{array}
$$

For each such p, $n(6, p) \leqq 4 \cdot 6$.

LEMMA 11. *If $p \equiv 1 \pmod{k}$ and if every integer is congruent modulo p to a sum of s k-th powers, then $n(k, p) \leqq sk$.*

There exist k roots h_i of $h^k \equiv 1$, whence $\Sigma h_i{}^j \equiv 0 \pmod{p}$ for $1 \leqq j < k$ by Newton's identities. We may take $f(x) = x^k + \cdots$. Then

$$\sum_{i=1}^{k} f(h_i y) \equiv k y^k \pmod{p}.$$

Since r is now k in Lemma 2, we have

LEMMA 12. *If $p \equiv 1 \pmod{k}$, $n(k, p) \leqq k^2$.*

For $k = 6$, $p = 31$, we find that $s = 4$ in Lemma 11, whence $n(6, 31) \leqq 24$. For $p = 7$, apply Lemma 9.

THEOREM 16. *If $p > 3$, $n(6, p) \leqq 24$.*

37. *Case $k = 8$.* Proof that $n(8, p) \leqq 72$.

I. $p \equiv -1 \pmod{4}$. By Lemma 6, there are exactly $p + 1$ solutions of

(175) $$h^8 + H^8 \equiv -1 \pmod{p}.$$

By Lemma 8, (175) has at most 48 solutions in common with one of $h^6 + H^6 \equiv -1$, $h^4 + H^4 \equiv -1$; $h^2 + H^2 \equiv -1$. Hence if $p + 1 > 48$ there is a solution of (175) for which $M_6 \not\equiv 0$, one for which $M_4 \not\equiv 0$, one for which $M_2 \not\equiv 0$, where the notations refer to (174) with $(j = 1, \cdots, 6)$, whence $n(8, p) \leqq 3n(6, p) \leqq 72$. For $p + 1 \leqq 48$, we have $n(8, p) \leqq 46$ by Lemma 9.

II. $p \equiv 1 \pmod{8}$. Then $n(8, p) \leqq 64$ by Lemma 12.

LEMMA 13. *Let* $p = 4f + 1 = x^2 + 4y^2$, $x \equiv 1$ (*mod* 4). *The number* N *of solutions of* $h^4 + H^4 \equiv -1$ (*mod* 4) *is* $p - 3 - 6x$ *if* f *is even, but* $p + 1 - 6x$ *if* f *is odd.*

For, by Theorems 2 and 5, $N = 8 + 16(00)$ or $16(02)$. Apply (52) and (56).

III. $p \equiv 5 \pmod 8$. If in Lemma 13, $p + 1 - 6x = 0$, eliminate p from $x^2 + 4y^2 = p$. Thus $z^2 + y^2 = 2$, $z = \frac{1}{2}(x - 3)$. Hence $z = \pm 1$, $x = 1$ or 5, $p = 5$ or 29.

Let $p \not\equiv 5$, $p \not\equiv 29$. Then $h^4 + H^4 \equiv -1$ has solutions. By Lemma 5 it has the same number of solutions as (175). Thus

$$h_1^8 + h_2^8 + h_3^8 \equiv 0 \pmod p \tag{176}$$

has solutions with h_3 prime to p. There exists an integer e belonging to the exponent 4 modulo p. We have (174) with $(j = 1, \cdots, 6)$. Let the new (174) be identically $\equiv 0 \pmod p$ for all solutions of (176). As below (174), $f(x)$ involves only even powers of x. We do not alter (176) if we replace h_3 by eh_3. If the old M_2 is $\equiv 0$, the new M_2 is $\not\equiv 0$, whence $c_2 \equiv 0$. Similarly $c_6 \equiv 0$. Hence $f(x) \equiv x^8 + c_4x^4$. Employ

$$4(a + b)^4 + 4(a - b)^4 + (2a)^4 + (2b)^4 = 24(a^2 + b^2)^2.$$

The corresponding sum of eighth powers is $8S$, where

$$S = 33(a^8 + b^8) + 28(a^6b^2 + a^2b^6) + 70a^4b^4.$$

We may take $a = 1$, $b^2 \equiv -1 \pmod p$. Then $S \equiv 80$. Hence there exist solutions of

$$\sum_{i=1}^{10} h_i^4 \equiv 0, \qquad \sum h_i^8 \equiv 8 \times 80 \not\equiv 0 \pmod p.$$

Then $f(h_iy) \equiv 640y^8$. In Lemma 2, r is now 4. Hence every integer is congruent to a sum of 4×10 values of $x^8 + c_4x^4$.

For $p = 29$, $n(8, p) \leqq 28$ by Lemma 9.

THEOREM 17. $n(8, p) \leqq 72$.

38. *Case* $k = 2q$, q *a prime* > 3.

I. $p \equiv 1 \pmod q$. Then $n(k, p) \leqq k^2$ by Lemma 12.

II. $p \not\equiv 1 \pmod q$. By Lemmas 6, 7, there are exactly $p \pm 1$ solutions of

$$h^k + H^k \equiv -1 \pmod p, \tag{177}$$

according as $p \equiv \mp 1 \pmod 4$. Apply Lemma 8 with $m < k$. Thus $n(k,p) \leqq 3n(k-2,p)$ if $p \pm 1 > k(k-2)$.

THEOREM 18. *If* $p \geqq 7$, $n(10,p) \leqq 216$.

This was proved if $p \pm 1 > 80$. But if $\leqq 79$, apply Lemma 9.

39. *Case* $k = 12$. To prove that $n(12,p) \leqq 648$.

I. $p \equiv 1 \pmod{12}$. By Lemma 12, $n(12,p) \leqq 144$.

II. $p \equiv -1 \pmod{12}$. By Lemma 6,

$$h^{12} + H^{12} \equiv -1 \pmod p \tag{178}$$

has exactly $p+1$ solutions. If $p+1 > 120$, Lemma 8 shows that $n(12,p) \leqq 3n(10,p) \leqq 648$. If $p+1 \leqq 120$, apply Lemma 9.

III. $p \equiv 5 \pmod{12}$. Since every integer is congruent to a cube (Lemma 4), the number N of solutions of (178) is the same as the number of solutions of

$$z^4 + w^4 \equiv -1 \pmod p, \tag{179}$$

which is true also by Lemma 5. Here $p \equiv 5$ or $17 \pmod{24}$.

IV. $p \equiv 5 \pmod{24}$. By Lemma 13, $N = p + 1 - 6x$, $x \equiv 1 \pmod 4$, $x^2 + 4y^2 = p$. Thus $N \equiv 0 \pmod{24}$. Since -2 is a quadratic non-residue of every prime $p \equiv 5 \pmod 8$, $z^4 \not\equiv w^4$ in (179). Also, $w^4 \equiv -1$ implies $w^8 \equiv 1$, $w^{p-1} \equiv 1$, $w^4 \equiv 1$, whence $z \not\equiv 0$, $w \not\equiv 0$ in (179). But $z^4 \equiv 1$ has four roots. Hence N is a multiple of $2 \times 4 \times 4$. Hence N is divisible by 96.

If $N > 12 \times 10$, Lemma 8 gives $n(12) \leqq 3n(10)$. It remains to treat $N = 0$, $N = 96$. By III, § 37, $N = 0$ requires that $p = 5$ or 29. If $N = 96$, eliminate $p = 6x + 95$ from $x^2 + 4y^2 = p$. Hence

$$26 = v^2 + y^2, \qquad v = \tfrac{1}{2}(x-3) = \text{odd}, \qquad v = \pm 5, \pm 1,$$

$p = 53, 101, 173 (125)$. Apply Lemma 9.

V. $p \equiv 17 \pmod{24}$. By Lemma 13, $N = p - 3 - 6x$. Hence $N \equiv 8 \pmod{24}$. Since there exists a number e belonging to the exponent 8, $e^4 \equiv -1 \pmod 4$ has four roots. Thus (179) has eight solutions with $z \equiv 0$ or $w \equiv 0$, and hence has $N - 8$ solutions z, w both prime to p. If the quadratic residue 2 is a residue of a fourth power, there are four solutions of $2z^4 \equiv -1$ and hence 16 solutions of (179) with $z^4 \equiv w^4$, whence $N - 24$ is a multiple of $2 \times 4 \times 4$. This with $N \equiv 8 \pmod{24}$ gives $N \equiv 56 \pmod{96}$. Next let $z^4 \not\equiv w^4$, then $N - 8$ is a multiple of 32. Hence $N \equiv 8 \pmod{96}$.

By Lemma 8, it remains to treat $N \leqq 120$, whence $N = 8$, 56 or 104. Eliminate p and write $v = \frac{1}{2}(x - 3) =$ odd.

If $N = 8$, $5 = v^2 + y^2$, $v = \pm 1$, $p = 17$ or 41.

If $N = 56$, $17 = v^2 + y^2$, $v = \pm 1$, $p = 89$, or $53 \not\equiv 17 \pmod{24}$.

If $N = 104$, $29 = v^2 + y^2$, $v = \pm 5$, $p = 65$ or 185 (not primes).

For these primes apply Lemma 9.

VI. $p \equiv 7 \pmod{12}$. By Lemma 5, the number N of solutions of (178) is the same as that of $x^6 + y^6 \equiv -1$. If $N = 0$, Lemma 10 gives $p \leqq 223$, whence Lemma 9 applies. Let $N > 0$. Since there exists an integer belonging to the exponent 6 modulo p, the usual proof gives $n(12) \leqq 3n(10)$ unless $f(x) = x^{12} + cx^6$.

Suppose that (178) has a solution in common with

$$h^6 + H^6 \equiv -1 \pmod{p}. \tag{180}$$

Elimination of h^6 gives $H^{12} + H^6 + 1 \equiv 0$, whence $H^6 \not\equiv 1$. Thus the g. c. d. of the exponents in $H^{18} \equiv 1$, $H^{p-1} \equiv 1$ must exceed 6. Hence $p \equiv 19 \pmod{36}$. Conversely, $H^{12} + H^6 + 1 \equiv 0$ then has twelve roots. Write $h = tH^2$, $t^6 \equiv 1$. Then (180) holds and there are exactly 72 simultaneous solutions of (178) and (180).

It therefore remains only to treat the case $N = 72$. By the results below Lemma 10, $N = p + 1 + 16A$ or $p + 1 + 2X$. Eliminate p. In the first case, $15 = z^2 + 3y^2$, $z = \frac{1}{3}(-A - 8)$, which is impossible. In the second case, $37 \cdot 12 = v^2 + 3w^2$. Thus $v = 3u$, $148 = 3u^2 + w^2$. The only solutions are $(u^2, w^2) = (9, 121), (16, 100), (49, 1)$. The ps are 19, 199, 271, 91 and 217 (factors 7), $73 \not\equiv 19 \pmod{36}$. Apply Lemma 8.

THEOREM 19. *If* $p \geqq 7$, $n(12, p) \leqq 648$.

40. *Case* $k = 14$. If $p \pm 1 > 14 \times 12 = 168$, § 38 applies. If $p \pm 1 \leqq 168$, Lemma 9 applies. Hence

THEOREM 20. *If* $p \geqq 7$, $n(14, p) \leqq 1944$.

41. *Case* $k = 16$. To prove $n(16, p) \leqq 5832$.

I. $p \equiv 1 \pmod{16}$. By Lemma 12, $n(16, p) \leqq 256$.

II. $p \equiv -1 \pmod{4}$. By Lemma 6, with $q = 1$, there exist exactly $p + 1$ solutions of

$$h^{16} + H^{16} \equiv -1 \pmod{p}. \tag{181}$$

By Lemma 8, $n(16, p) \leqq 3n(14, p)$ unless $p + 1 \leqq 224$, and then Lemma 9 applies.

III. $p \equiv 5 \pmod 8$. By Lemma 5, (181) has the same number N of solutions as $x^4 + y^4 \equiv -1 \pmod p$. By Lemma 13,

$$N = p + 1 - 6x, \qquad p = x^2 + 4y^2, \qquad x \equiv 1 \pmod 4.$$

Since -2 is a quadratic non-residue of p, $h^{16} \not\equiv H^{16}$ in (181). Also $h \not\equiv 0$, $H \not\equiv 0$. If $j^{16} \equiv 1$, then $j^4 \equiv 1$. Hence the solutions fall into sets of $2 \times 4 \times 4$, so that N is a multiple of 32. There remains the case $N \leqq 224$. If $N \equiv 1 \pmod 3$, p would be divisible by 3.

If $N = 0$, $p = 5$ or 29 by III of § 37. If $N = 96$, $p = 53$, 101 or 173 by IV of § 39. Write $v = \frac{1}{2}(x - 3) =$ odd. If $N = 32$, $10 = v^2 + y^2$, $p = 13$, 37, 61. If $N = 128$, $34 = v^2 + y^2$, $p = 109$ or 181. If $N = 192$, $50 = v^2 + y^2$, $v^2 = 1, 25, 49$, $p = 149$, 197, 269, 293. If $N = 224$, $58 = v^2 + y^2$, $p = 157$ or 277. For these p's apply Lemma 9.

IV. $p \equiv 9 \pmod{16}$. By Lemma 5, (181) has the same number M of solutions as

$$h^8 + H^8 \equiv -1 \pmod p. \tag{182}$$

By Theorem 5, $M = 64(04)$. By (114), (115),

$$M = p + 1 - 18x \text{ or } M = p + 1 + 6x + 24a, \quad p = x^2 + 4y^2, \quad p = a^2 + 2b^2. \tag{183}$$

First, let $M > 0$. Since there exists an integer belonging to the exponent 8, $n(16) \leqq 3n(14)$ unless $f(x) = x^{16} + cx^8$. As in VI, § 39, if there be a simultaneous solution of (181) and (182), then

$$H^{16} + H^8 + 1 \equiv 0, \quad H^{24} \equiv 1, \quad H^8 \not\equiv -1 \pmod p, \quad p \equiv 25 \pmod{48}.$$

Conversely, there are then exactly $16 \times 8 = 128$ common solutions. Hence there remains only the case $M = 128$. Then in (183_1), $p = 18x + 127$, whence 52 is the sum of the squares of $\frac{1}{2}(x - 9)$ and y, viz., 36 and 16, or vice versa. Thus $p = 73$ or $433 \equiv 1 \pmod{16}$. Next, if $M = 128$ in (183_2), then

$$p + 1 - 30\sqrt{p} < 128, \qquad p < 1156.$$

But for $p = 73$ or $p < 1156$, Lemma 9 applies.

Second, let $M = 0$. Evidently $p < 1156$.

THEOREM 21. *If* $p > 7$, $n(16, p) \leqq 5832$.

42. *Case* $k = 18$. Let $p \equiv -1 \pmod 3$. By Lemma 6, with $q = 9$,

$$(184) \qquad h^{18} + H^{18} \equiv -1 \pmod{p}$$

has exactly $p+1$ solutions. By Lemma 8, if $p+1 > 288, n(18, p) \leqq 3n(16, p)$. Apply Lemma 9.

For $p > 2$, there remains the case $p \equiv 1 \pmod{6}$. If $p \equiv 1 \pmod{18}$, apply Lemma 12. Henceforth, let $p \equiv 7$ or $13 \pmod{18}$. By Lemma 5, (184) has the same number N of solutions as $x^6 + y^6 \equiv -1 \pmod{p}$. By Lemma 8, there remains the case $N \leqq 18 \cdot 16 = 288$.

In the respective cases below Lemma 10,

$$p + 1 \leqq -16A + 288 < 16\sqrt{p} + 288, \qquad p < 729,$$

$$p + 1 \leqq 288 - 2X \leqq 288 + 2\sqrt{37p}, \qquad p \leqq 580.$$

Apply Lemma 9. Hence we have

THEOREM 22. *If* $p \geqq 7$, $n(18, p) \leqq 17496$.

43. *Case* $k = 20$. When $p \not\equiv 1 \pmod{5}$, $p \equiv -1$ or $+1 \pmod{4}$, we find by Lemma 6 or Lemma 13 the number N of solutions of $h^{20} + H^{20} \equiv -1 \pmod{p}$, and proceed as usual. If $p \equiv 1 \pmod{20}$, apply Lemma 12. There remains only the case $p \equiv 11 \pmod{20}$; since N is not known, we resort to the rough Theorem 15 and Theorem 14 and obtain

THEOREM 23. $n(20) \leqq 11\, n(19) \leqq 22\, n(18)$.

44. For $k = 22$, we employ § 38 with $q = 11$. It remains to treat $p \pm 1 \leqq 440$; apply Lemma 9.

THEOREM 24. $n(22, p) \leqq 3\, n(20, p)$.

45. It remains to treat primes p which divide k. Let p^t be the highest power of p which divides k. Write

$$P = p^{t+1} \text{ if } p > 2, \qquad P = p^{t+2} \text{ if } p = 2.$$

We seek N such that every integer is congruent modulo P to a sum of N values of any polynomial in x not all of whose values are multiples of p. By Lemma 9, $N \leqq P - 1$. Hence if $k = 3$, $N \leqq 8$; if $k = 4$, $N \leqq 15$; *if* $k = 5$, $N \leqq 24$; *if* $k = 6$, $N \leqq 8$. For these, $N \leqq s$ in (160). For $7 \leqq k \leqq 26$, $N < n = n(k)$, for the n listed in § 29.

THE UNIVERSITY OF CHICAGO.

THE EQUIVALENCE OF NON-SINGULAR PENCILS OF HERMITIAN MATRICES IN AN ARBITRARY FIELD.

By J. Williamson.

The problem of the equivalence of two non-singular pencils of real symmetric matrices in the real field was first solved by Muth.[1] More recently Trott,[2] Wegner,[3] Ingraham[4] and Turnbull[5] have solved the similar problem for two Hermitian matrices under conjunctive transformations in the complex field. The notation used by Trott was such, that he was able to discuss the Hermitian case and at the same time the real symmetric case. In this paper we show how Trott's method may be extended to the similar problem of the equivalence of two non-singular pencils of Hermitian (or symmetric) matrices with respect to a general commutative field K. Incidentally, as is often the case with a generalization, we show why the results in the case of the complex field (or real field) are comparatively simple. We prove that a necessary and sufficient condition for two such pencils to be equivalent is that;

(α) *they have the same elementary factors with respect to* K,
and (β) *certain diagonal matrices be equivalent in over fields of* K.

In the simple cases already considered conditions (β) can all be expressed in terms of the equality of certain integers—the signatures of the respective quadratic or hermitian forms. That no such great simplification is possible in the general case is apparent from a consideration of two pairs of one rowed matrices a, b, and c, d in the rational field, where a, b, c, d, are all rational numbers and b and d are both different from zero. The pair a, b, is equivalent to the pair c, d, if, and only if, $a - \lambda b$ and $c - \lambda d$ have the same elementary

[1] P. Muth, "Über reele Äquivalennz von Scharen reeler quadratischer Formen," *Crelle's Journal*, vol. 128 (1905), pp. 302-343.

[2] G. R. Trott, "On the canonical form of a non-singular pencil of Hermitian matrices," *American Journal of Mathematics*, vol. 56, no. 3 (1934), pp. 359-371. We shall refer to this paper as Trott, 1.

[3] K. W. Wegner, "Equivalence of pairs of Hermitian matrices," *Bulletin of the American Mathematical Society*, vol. 40, no. 1, January (1934), Abstract 103.

[4] M. H. Ingraham, "The singular case of the equivalence of pairs of Hermitian matrices," *Bulletin of the American Mathematical Society*, vol. 40, no. 7, July (1934), Abstract 242.

[5] H. W. Turnbull, "Pencils of Hermitian forms," *Proceedings of the London Mathematical Society*, series 2, vol. 39 (1935), pp. 232-248.

divisors, i. e., if $a/b = c/d$ (condition α), and, if $b = k^2 d$, where k is a rational number (condition β).

Section I is devoted to preliminary definitions and proofs; the main results are proved in § (2) and a short discussion of these results is given in § (3). No attempt is made to consider a similar problem for singular pencils.

1. Let K be any commutative field of characteristic zero[6] and let $K(i)$ be a quadratic field over K, where i is a root of the equation $x^2 - \alpha = 0$, irreducible in K. Then every element a of $K(i)$ is of the form $a = a_1 + ia_2$, where a_1 and a_2 lie in K, so that the conjugate of a is the element $\bar{a} = a_1 - ia_2$. If R is a matrix over $K(i)$, $R = R_1 + iR_2$, where R_1 and R_2 are both matrices over K, and $\bar{R} = R_1 - iR_2$. The matrix R^* is defined to be the conjugate transposed of R so that

$$R^* = \bar{R}' = R'_1 - iR'_2.$$

When R is a square matrix of order n, we may consider R as a matrix of matrices and write

$$(1) \qquad R = (R_{ij}), \qquad (i, j = 1, 2, \cdots, t),$$

where R_{ij} is a matrix of r_i rows and r_j columns and $r_1 + r_2 + \cdots r_t = n$. If S is a second n-rowed square matrix and S is written as a matrix of matrices,

$$(2) \qquad S = (S_{ij}), \qquad (i, j = 1, 2, \cdots, t),$$

where S_{ij} is also a matrix of r_i rows and r_j columns, we say that S and R are *similarly partitioned* or that (2) is a partition of S similar to (1). If in (1), when i is different from j, R_{ij} is the zero matrix, we call R a *diagonal block matrix* and write

$$R = [R_{11}, R_{22}, \cdots, R_{tt}].$$

If D is a square matrix of order n, whose elements lie in K, the invariant factors $E_j(\lambda)$ of $D - \lambda E$ are polynomials over K. We call[7] the powers of the distinct irreducible factors of $E_j(\lambda)$ the elementary factors (with respect to K) of $D - \lambda E$. Let the elementary factors of $D - \lambda E$ be

$$(3) \qquad [p_i(\lambda)]^{\eta_{ij}},$$
$$(i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, k_i;\ \eta_{ij} \geqq \eta_{is} \geqq 1, \text{ if } j < s),$$

[6] It is not essential for this discussion that the characteristic p of K be zero. On the other hand p cannot be arbitrary. We, however, restrict ourselves to the case $p = 0$ for the sake of simplicity.

[7] Cf. Neal McCoy, "On the rational canonical form of a function of a matrix," *American Journal of Mathematics*, this volume, p. 492; J. H. M. Wedderburn, *Lectures on Matrices*, pp. 123-126.

where $p_i(\lambda)$ is a polynomial over K of degree n_i, irreducible in K, with leading coefficient unity and such that $p_i(\lambda) \neq p_j(\lambda)$, if $i \neq j$. Then $n = \sum_{i=1}^{t} n_i \sum_{j=1}^{k_i} \eta_{ij}$. Further let p_i be a square matrix of order n_i, with elements in K, whose characteristic equation is $p_i(\lambda) = 0$, and let N_{ij} be the matrix

$$(4) \qquad N_{ij} = \begin{bmatrix} p_i & e_i & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & p_i & e_i & \cdot & \cdot & \cdot & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & e_i \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & p_i \end{bmatrix} \qquad \begin{matrix}(i = 1, 2, \cdots, t; \\ j = 1, 2, \cdots, k_i),\end{matrix}$$

where e_i is the unit matrix of order n_i and N_{ij}, considered as a matrix of matrices, is of order η_{ij}. If M_i is the diagonal block matrix

$$(5) \qquad M_i = [N_{i1}, N_{i2}, \cdots, N_{ik_i}], \qquad (i = 1, 2, \cdots, t),$$

and

$$(6) \qquad M = [M_1, M_2, \cdots, M_t],$$

the elementary factors of $D - \lambda E$ are the same as those of $M - \lambda E$. Hence M is similar to D in K and is a canonical form of D in K.

We now define two matrices of order η_{ij}, whose elements are matrices of order n_i. These two matrices are the auxiliary unit matrix

$$(7) \qquad U_{ij} = \begin{bmatrix} 0 & e_i & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & e_i & \cdot & \cdot & \cdot & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & e_i \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}$$

and the counter unit matrix

$$(8) \qquad T_{ij} = \begin{bmatrix} 0 & \cdot & \cdot & \cdot & 0 & e_i \\ 0 & \cdot & \cdot & \cdot & e_i & 0 \\ . & . & . & . & . & . \\ e_i & \cdot & \cdot & \cdot & 0 & 0 \end{bmatrix}.$$

The matrix N_{ij}, defined by (4), may accordingly be written in the convenient form

$$N_{ij} = p_i E_{ij} + U_{ij},$$

where $E_{ij} = T^2_{ij}$. Moreover a simple calculation shows that

$$(9) \qquad U'_{ij} T_{ij} = T_{ij} U_{ij}.$$

Let q_i be a non-singular matrix over K of order n_i, satisfying the equation

$$(10) \qquad q_i p_i = p'_i q_i, \qquad (i = 1, 2, \cdots, t).$$

It has been shown that such a matrix q_i exists and that it is necessarily symmetric.[8] Further, if,

$$(11) \qquad Q_{ij} = q_i T_{ij}, \qquad (i = 1, 2, \cdots, t; \; j = 1, 2, \cdots, k_i),$$

then

$$\begin{aligned} q_i T_{ij} N_{ij} &= q_i T_{ij}(p_i E_{ij} + U_{ij}), \\ &= (p'_i E_{ij} + U'_{ij}) q_i T_{ij}, \text{ by (9) and (10),} \end{aligned}$$

so that

$$(12) \qquad Q_{ij} N_{ij} = N'_{ij} Q_{ij}.$$

Accordingly the matrix

$$(13) \qquad Q = [Q_1, Q_2, \cdots, Q_t], \quad \text{where} \quad Q_i = [Q_{i1}, Q_{i2}, \cdots, Q_{ik_i}],$$

is a non-singular symmetric matrix over K, satisfying the equation

$$(14) \qquad QM = M'Q.$$

Moreover, if R is a matrix over $K(i)$, such that

$$(15) \qquad RM = M'R,$$

then

$$(16) \qquad R = QS,$$

where S is a matrix commutative with M. The form of S is known.[9] In fact S is a diagonal block matrix

$$(17) \qquad S = [S_1, S_2, \cdots, S_t]$$

partitioned similarly to M in (6). Further, if for simplicity we write η_j, T_j, U_j, q, p, and k for η_{ij}, T_{ij}, U_{ij}, q_i, p_i and k_i respectively and let

$$(18) \qquad S_i = (S_{rs}), \qquad (r, s = 1, 2, \cdots, k),$$

be a partition of S_i similar to that of M_i in (5), S_{rs} is a matrix of η_r rows and η_s columns, where $\eta_r \geqq \eta_s$, if $r \leqq s$. Moreover, if $r \leqq s$,

$$(19) \qquad S_{rs} = \begin{pmatrix} G_{rs} \\ 0 \end{pmatrix}, \qquad S_{sr} = (0' G_{sr}),$$

---

[8] R. C. Trott, *Bulletin of the American Mathematical Society*, vol. 41, no. 1, part 2, January (1935), Abstract No. 95. We shall refer to this paper as Trott 2.

[9] Trott 2.

where G_{rs} and G_{sr} are both square matrices of order $\eta_s = \eta$, while 0 denotes the zero matrix of orders $\eta_r - \eta, \eta$ and $0'$ its transposed. More exactly,

$$G_{rs} = \sum_{a=0}^{\eta-1} g_{rsa} U_s^a, \qquad G_{sr} = \sum_{a=0}^{\eta-1} g_{sra} U_s^a, \tag{20}$$

where g_{rsa} and g_{sra} are polynomials in the matrix p with coefficients in $K(i)$.

We now define the two matrices

$$\bar{S}_{rs} = (0, \bar{G}_{rs}), \quad \bar{S}_{sr} = \begin{pmatrix} \bar{G}_{sr} \\ 0 \end{pmatrix} \qquad r \leqq s, \tag{21}$$

so that in particular, if $\eta_r = \eta_s$,

$$\bar{S}_{rs} = \bar{S}_{rs}, \qquad \eta_r = \eta_s. \tag{22}$$

It should be noted that $\bar{S}_{rs}$ is formally the transposed conjugate of S_{rs}, if p is considered as an indeterminate instead of a matrix.

It follows from (20) that

$$\begin{aligned} G^*_{rs} q T_s &= \sum_{a=0}^{\eta-1} g^*_{rsa} U'^a_s q T_s, \\ &= q T_s \sum_{a=0}^{\eta-1} \bar{g}_{rsa} U_s^a \text{ by (9) and (10)}, \\ &= q T_s \bar{G}_{rs}. \end{aligned}$$

Hence, if $r \leqq s$,

$$\begin{aligned} S^*_{rs} q T_r &= (G^*_{rs}\ 0) q T_r, \text{(0 the zero matrix of orders } \eta_r \rightarrow \eta_s, \eta_s) \\ &= (0\ G^*_{rs} q T_s) = (0\ q T_s \bar{G}_{rs}) = q T_s (0\ \bar{G}_{rs}) = q T_s \bar{S}_{rs} \text{ by (21)}. \end{aligned}$$

Similarly,

$$S^*_{sr} q T_s = \begin{pmatrix} 0 \\ G^*_{sr} \end{pmatrix} q T_s = \begin{pmatrix} 0 \\ q T_s \bar{G}_{sr} \end{pmatrix} = q T_r \begin{pmatrix} \bar{G}_{sr} \\ 0 \end{pmatrix} = q T_r \bar{S}_{sr}.$$

Therefore for all values of r and s

$$S^*_{rs} q T_r = q T_s \bar{S}_{rs}. \tag{23}$$

If the matrix R, defined by (16), is such that $R = R^*$, on equating corresponding elements of the two matrices we have

$$q T_r S_{rs} = (q T_s S_{sr})^*,$$

or

$$q T_r S_{rs} = S^*_{sr} q T_s = q T_r (\bar{S}_{sr}) \text{ by (23)}, \tag{24}$$

so that

$$S_{rs} = \bar{S}_{sr}. \tag{25}$$

In particular, if $\eta_r = \eta_s$, it follows from (22) that

(26) $$S_{rs} = \bar{S}_{sr}, \qquad (r, s = 1, 2, \cdots, k)$$

and, if $r = s$, that

(27) $$S_{rr} = \bar{S}_{rr}, \qquad (r = 1, 2, \cdots, k),$$

so that S_{rr} lies in K.

2. We now consider two square matrices, A and B, of order n, with elements in the field $K(i)$, of which the second, B, is non-singular. The matrices A and B are such that $A = A^*$ and $B = B^*$, so that both matrices are hermitian matrices or else, when A and B are both matrices over K, symmetric matrices over K. Moreover, if A and B are both matrices over K, in the sequel every matrix P is a matrix over K and P^* is to be interpreted as P'. Since $A = A^*$ and $B = B^*$, the invariant factors $E_j(\lambda)$ of the pencil $A - \lambda B$, which are certainly polynomials over the field $K(i)$, are unaltered by the substitution of $-i$ for i and are accordingly polynomials over K. We are therefore at liberty to talk of the elementary factors (with respect to K) of the pencil $A - \lambda B$. We let these elementary factors be the polynomials (3), so that $A - \lambda B$ has the same elementary factors as $M - \lambda E$. Since the elementary factors of $A - \lambda B$ are the same as those of $AB^{-1} - \lambda E$, the two matrices AB^{-1} and M are similar. Hence there exists a non-singular matrix P, such that

$$P^{-1}(AB - \lambda E)P = M - \lambda E,$$

or

(28) $$(A - \lambda B)B^{-1}P = P(M - \lambda E).$$

In general the elements of the matrix P lie in $K(i)$, but, if A and B are both symmetric matrices over K, P is also a matrix over K. As a consequence of (28) we have

(29) $$P^*B^{-1}(A - \lambda B)B^{-1}P = R(M - \lambda E),$$

where

(30) $$R = P^*B^{-1}P.$$

It follows from (30), since $B^* = B$, that $R = R^*$ and from (29) that $RM = P^*B^{-1}AB^{-1}P$, so that RM is hermitian, and accordingly that

(31) $$RM = M^*R^* = M'R.$$

We shall now reduce the pencil of matrices $R(M - \lambda E)$ by a *conjunctive*

transformation[10] to a canonical form $G(M-\lambda E)$; that is, determine a non-singular matrix W such that

$$(32) \qquad W^*R(M-\lambda E)W = G(M-\lambda E).$$

Accordingly, as a consequence of (29) and (32), the pencil $A-\lambda B$ is equivalent under a conjunctive transformation to the pencil $G(M-\lambda E)$. Moreover, it follows immediately from (32) that $G=W^*RW$ and that $W^*RMW = GM = W^*RWM$. Hence $G=G^*$ and

$$(33) \qquad MW = WM,$$

so that throughout the various stages of the reduction the transforming matrices are all permutable with M.

As a consequence of (31) $R=QS$, where Q is defined by (13) and S by (17), (18) and (19). Therefore,

$$R = [R_1, R_2, \cdots, R_t],$$

is a diagonal block matrix, where

$$R_i = Q_i S_i, \qquad (i=1,2,\cdots,t),$$

and, since $MW=WM$, the matrix W is also a diagonal block matrix $[W_1, W_2, \cdots, W_t]$, where W_i is of the same order as M_i. Hence we see that in reducing R we may reduce each R_i separately by transformations W_i permutable with M_i. As this is the case we temporarily drop all suffixes i and write M, R, S, q, T_j etc. for $M_i, R_i, S_i, Q_i, T_{ij}$ respectively.

We first show that without any loss of generality we may assume S_{11} to be non-singular. Since $\eta_1 \geqq \eta_2 \geqq \cdots \geqq \eta_k \geqq 1$, we may suppose that $\eta_1 = \eta_2 = \cdots = \eta_s > \eta_{s+1}$, $1 \leqq s \leqq k$. If S_{11}, is singular but, for some value of $j \leqq s$, S_{jj} is non-singular, by interchanging the first row and the j-th row of S and the first column and the j-th column, we move S_{jj} into the place of S_{11}. Moreover such an interchange may be accomplished by means of a conjunctive transformation permutable with M and Q.[11] We now assume that S_{jj} is singular for all values of j, $1 \leqq j \leqq s$. If s_{ij} denotes the first element, i. e., the element in the first row and first column, of the matrix S_{ij}, $|S_{jj}| = |s_{jj}|^{\eta_j}$ and, since s_{jj} is a polynomial in the matrix p with coefficients

[10] We shall use the term conjunctive transformation to include the case of a congruent transformation; i. e., a transformation of matrix W, where W lies in K, so that $W^* = W'$.

[11] Turnbull and Aitken, *An Introduction to the Theory of Canonical Matrices*, p. 11.

in K (equation (27)), $s_{jj}=0$. In particular $s_{11}=0$, and, since, from the nature of S_{i1} (equation (19)), $s_{i1}=0$ when $i>s$, there is at least one value j, $1<j\leqq s$, such that $s_{j1}\neq 0$, as otherwise S would be singular. After a suitable interchange of rows and columns we may therefore suppose that s_{21} is not zero. Let

$$W_1=\left[\begin{pmatrix} E_1 & 0 \\ E_1 & E_1 \end{pmatrix},\ E_2\right] \quad \text{and} \quad W_2=\left[\begin{pmatrix} E_1 & 0, \\ iE_1 & E_1 \end{pmatrix},\ E_2\right],$$

where E_1 and E_2 are the unit matrices of orders $n_i\eta_1$ and $n_i(\eta_3+\eta_4+\cdots+\eta_m)$ respectively. The two matrices W_1 and W_2 are both permutable with M as are the matrices W^*_1 and W^*_2 with Q. A simple calculation shows that, if

$$W^*_1QSW_1=QX \quad \text{and} \quad W^*_2QSW=QY$$

and X and Y are partitioned similarly to M,

$$X_{11}=S_{11}+S_{21}+S_{12}+S_{22},\quad Y_{11}=S_{11}+i(S_{12}-S_{21})-i^2S_{22}.$$

The first two elements of these matrices are respectively $x_{11}=s_{12}+s_{21}$ and $y_{11}=i(s_{12}-s_{21})$, since $s_{11}=s_{22}=0$. As $s_{21}\neq 0$, at least one of x_{11} or y_{11} is different from zero, so that at least one of X_{11} or Y_{11} is non-singular. However, as W_2 is not a matrix over K, we must still show that, if S is a matrix over K, the matrix X_{11} is non-singular. This is in fact the case; for, since $S^*=S$, by (26) $s_{21}=\bar{s}_{12}$, so that if S lies in K, $s_{21}=\bar{s}_{12}=s_{12}$ and $x_{11}=2s_{21}\neq 0$. Hence we may assume without any loss of generality that S_{11} is non-singular.

We next show that S may be reduced to a diagonal block matrix partitioned similarly to M. Let

$$W=\begin{bmatrix} E_1 & -S_{11}^{-1}S_{12} & -S_{11}^{-1}S_{13} & \cdots & -S_{11}^{-1}S_{1k} \\ 0 & E_2 & 0 & \cdots & 0 \\ 0 & 0 & E_3 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & E_k \end{bmatrix},$$

so that W is certainly permutable with M. The element in the r-th place, $r>1$, of the first column of W^*Q is

$$-S^*_{1r}(S^*_{11})^{-1}qT_1=-S^*_{1r}qT_1S_{11}^{-1}=-qT_rS_{r1}S_{11}^{-1} \text{ by (24).}$$

Hence $W^*Q=QH$, where H is obtained from W^* by replacing the element in the r-th place, $r>1$, of the first column by $-S_{r1}S_{11}^{-1}$. A simple calculation now shows that if,

$$W^*QSW = QD = Q(D_{rs}), \qquad (r, s = 1, 2, \cdots, k),$$
$$D_{11} = S_{11}, \quad D_{r1} = D_{1r} = 0, \qquad (r \neq 1),$$
$$D_{rs} = S_{rs} - S_{r1}S_{11}^{-1}S_{1s}, \qquad (r, s = 2, 3, \cdots, k).$$

We have therefore shown that there exists a non-singular conjunctive transformation permutable with M, which reduces $R = QS$ to the form $Q[S_{11}, H]$, where H is a square matrix of order $n_i(\eta_2 + \eta_3 + \cdots \eta_k)$ and is commutative with $[N_2, N_3, \cdots, N_k]$. Accordingly H is of exactly the same type as S and we may repeat our previous argument with S replaced by H. Hence in $k - 1$ steps we deduce the existence of a non-singular matrix V, permutable with M, such that

$$V^*QSV = Q[G_1, G_2, \cdots, G_k], \tag{34}$$

where G_j is of the same order as N_j, $(j = 1, 2, \cdots, k)$. Moreover G_j is permutable with N_j and Q_jG_j is hermitian. Hence by (27) G_j is a matrix over K and Q_jG_j is symmetric.

We now show that it is possible to reduce G_j, $(j = 1, 2, \cdots, k)$, to a diagonal block matrix. For simplicity we write $\eta = \eta_j$, so that

$$G_j = \sum_{a=0}^{\eta-1} g_a U_j^a \quad \text{(formulas (19) and (20))},$$

where g_a is a polynomial in p with coefficients in K. Since G_j is non-singular, g_0 is non-singular and is accordingly different from zero. If $g_1 = g_2 = \cdots = g_c = 0$ and $c = \eta - 1$, G is a diagonal block matrix. If $c < \eta - 1$ and $g_{c+1} \neq 0$, we consider the matrix

$$W = E_j - wU_j^{c+1}, \text{ where } w = g_{c+1}/2g_0,$$

so that

$$G_jW^2 = G_j(E_j - 2wU_j^{c+1} + w^2U_j^{2c+2}),$$
$$= g_0E_j + h_{c+1}U_j^{c+1} + \sum_{a=c+2}^{\eta-1} h_aU_j^a = H_j,$$

where $h_{c+1} = g_{c+1} - 2gw = 0$ and h_a is a polynomial in p, when $a \geqq c + 2$. But

$$W^*Q_jG_jW = W'qT_jG_jW = qT_jWG_jW \text{ by (9) and (10)},$$
$$= qT_jG_jW^2 = qT_jH_j.$$

Hence the matrix W reduces Q_jG_j to Q_jH_j, where H_j is of the same form as G_j except that the coefficient of U_j^{c+1} is now zero. If the coefficient of U_j^{c+2} in H_j is different from zero, we may repeat our argument with G_j replaced

by H_j. Accordingly in at most $\eta - 1$ such steps we can reduce G_j to the diagonal form g_0E_j. Let us therefore suppose that G_j is already in diagonal form, so that, after an obvious change of notation,

$$(35) \qquad G_j = g_jE_j, \qquad (j = 1, 2, \cdots, k),$$

and $g_j = g_j(p)$ is a polynomial in p with coefficients in K. Equation (34) therefore becomes,

$$(36) \qquad V^*QSV = QG = Q[g_1E_1, g_2E_2, \cdots, g_kE_k],$$

and we call the matrix on the right of this last equation a *canonical form* for the matrix QS.

It is apparent that this canonical form may not be unique. Suppose therefore that there exists a non-singular matrix Y, permutable with M, such that

$$(37) \qquad Y^*QSY = QF = Q[f_1E_1, f_2E_2, \cdots, f_kE_k],$$

where $f_j = f_j(p)$ is a polynomial in p with coefficients in K. Then, if $W = V^{-1}Y$, W is permutable with M and, as a consequence of (36) and (37),

$$(38) \qquad W^*QGW = QF.$$

If $W = (W_{rs})$, $(r, s = 1, 2, \cdots, k)$, is a partition of W similar to that of M, W_{rs} and W_{sr} are of the same forms as S_{rs} and S_{sr} in (19). We define $\bar{W}_{rs}$ in an analogous manner to $\bar{S}_{rs}$ in (21), so that in particular, if $\eta_r = \eta_s$, $\bar{W}_{rs} = \bar{W}_{rs}$. The matrix equation (38) may now be written in the form

$$\sum_{a=1}^{k} W^*_{ar}qT_ag_aW_{as} = \delta_{rs}qT_rf_rE_r, \quad (r, s = 1, 2, \cdots, k;\ \delta_{rs} \text{ the Kronecker } \delta).$$

Hence by (23)

$$qT_r \sum_{a=1}^{k} \bar{W}_{ar}g_aW_{as} = \delta_{rs}qT_rf_rE_r,$$

or, on dividing by the non-singular matrix qT_r,

$$(39) \qquad \sum_{a=1}^{k} \bar{W}_{ar}g_aW_{as} = \delta_{rs}f_rE_r.$$

If w_{ij} is the first element of the matrix W_{ij} and $\bar{w}_{ij}$ the first element of the matrix $\bar{W}_{ij}$, it follows from the nature of the matrices W_{ij} and $\bar{W}_{ij}$ (cf. equations (19) and (21)), that the first element of the matrix $\bar{W}_{ar}g_aW_{as}$ is $\bar{w}_{ar}g_aw_{as}$. Accordingly by equating the first elements of each component matrix in (39), we have

$$\sum_{\alpha=1}^{k} \bar{w}_{\alpha r} g_\alpha w_{\alpha s} = \delta_{rs} f_r. \tag{40}$$

But by (19) and (21),

$$w_{\alpha s} = 0, \text{ if } \eta_\alpha < \eta_s; \ \bar{w}_{\alpha r} = 0, \text{ if } \eta_\alpha > \eta_r; \ \bar{w}_{\alpha r} = \bar{w}_{\alpha r}, \text{ if } \eta_\alpha = \eta_r.$$

Hence, if $\eta_{c-1} > \eta_c = \eta_{c+1} = \cdots = \eta_d > \eta_{d+1}$ and $c \leqq s \leqq d$, $c \leqq r \leqq d$, $\bar{w}_{\alpha r} g_\alpha w_{\alpha s} = 0$, when $\alpha < c$ or $\alpha > d$. Accordingly (40) becomes

$$\sum_{\alpha=c}^{d} \bar{w}_{\alpha r} g_\alpha w_{\alpha s} = \delta_{rs} f_r. \tag{41}$$

Let D be the matrix (d_{ij}), $(i, j = 1, 2, \cdots, d - c + 1)$, where $d_{ij} = w_{c+i-1, c+j-1}$. Then it is a consequence of the form of W that, since W is non-singular, D is non-singular, for, after a proper interchange of rows and columns, it can be shown that $|D|$ is a factor of $|W|$. Each element d_{ij} of D is a polynomial $d_{ij}(p)$ in the matrix p with coefficients in $K(i)$ and therefore (41) may be written in the form of a matrix equation

$$\bar{D}[g_c, g_{c+1}, \cdots, g_d] D = [f_c, f_{c+1}, \cdots, f_d], \tag{42}$$

where $\bar{D} = (\bar{d}_{ij}) = (\bar{d}_{ji})$, $i, j = 1, 2, \cdots, d + 1 - c$; cf. equation (21). It is important to notice that $\bar{D}$ is not the same as D^*, since $D^* = (d^*_{ij})$, where $d^*_{ij} = \bar{d}_{ji}(p')$. Let x be an indeterminate and let $D(x)$ denote the matrix whose typical element is $d_{ij}(x)$. Then, if

(43) $|D(x)| = \rho(x) + i\sigma(x), \rho(x), \sigma(x)$ polynomials with coefficients in K,

(44) $\quad |D| = |\rho(p) + i\sigma(p)|$.[12]

Similarly $|\bar{D}| = |\rho(p) - i\sigma(p)|$, so that

$$|D| \ |\bar{D}| = |(\rho(p))^2 - i^2(\sigma(p))^2| = |\mu(p)|, \tag{45}$$

where $\mu(x)$ is a polynomial in x with coefficients in K. Since D, and similarly, $\bar{D}$, are both non-singular, $D\bar{D}$ is non-singular, so that by (45), $\mu(p)$ is non-singular. Hence, since $\mu(p)$ is a polynomial over K,

$$\mu(p) \neq 0, \tag{46}$$

is a necessary and sufficient condition that D and $\bar{D}$ both be non-singular.

If θ is a root of the irreducible equation $p(x) = 0$, the field $K(\theta)$ is simply isomorphic with the field formed by all polynomials in p with coefficients in K. Consequently it follows from (45) and (46) that

---

[12] J. Williamson, "The latent roots of a matrix of special type," *Bulletin of the American Mathematical Society*, vol. 37 (August, 1931), p. 587.

$$(47)\qquad |D(\theta)|\ |\widetilde{D(\theta)}| = \mu(\theta) \neq 0$$

and accordingly that both matrices $D(\theta)$ and $\widetilde{D(\theta)}$ are non-singular. Since the elements of $D(\theta)$ are no longer matrices, $\tilde{D}(\theta) = D^*(\theta)$, and we therefore have, as a consequence of (42),

$$(48)\qquad D^*(\theta)[g_c(\theta),\cdots,g_d(\theta)]D(\theta) = [f_c(\theta),\cdots,f_d(\theta)],$$

where $D(\theta)$ and $D^*(\theta)$ are both non-singular. In other words the two matrices $[g_c(\theta),\cdots,g_d(\theta)]$ and $[f_c(\theta),\cdots,f_d(\theta)]$ are conjunctively equivalent. Conversely, if (48) is true and both $D(\theta)$ and $D^*(\theta)$ are non-singular,[13] $\mu(\theta) \neq 0$ by (47) and accordingly (46) is satisfied, so that (42) is true, where D is non-singular. Hence not only does (42) imply (48) but also (48) implies (42).

Before summing up and stating our results in the form of a theorem it will prove convenient to alter our notation slightly. Accordingly we relabel the integers η_i in the following manner;

$$(49)\qquad \eta_1 = \eta_2 = \cdots = \eta_{s_1} = \xi_1 > \eta_{s_1+1} = \eta_{s_1+2} = \cdots = \eta_{s_1+s_2} = \xi_2 > \eta_{s_1+s_2+1} = \cdots \xi_{r-1} > \eta_{s_1+s_2+\ldots s_{r-1}+1} = \cdots = \eta_{s_1+s_2+\ldots+s_r} = \xi_r,$$

where $s_1 + s_2 + \cdots + s_r = k$, and write

$$(50)\qquad \begin{aligned} I_j &= [E_c, E_{c+1},\cdots,E_d], & L_j &= [N_c, N_{c+1},\cdots,N_d],\\ \gamma_j &= [g_c, g_{c+1},\cdots,g_d], & \phi_j &= [f_c, f_{c+1},\cdots,f_d],\\ \Gamma_j &= [g_cE_c, g_{c+1}E_{c+1},\cdots,g_dE_d], & \Phi_j &= [f_cE_c, f_{c+1}E_{c+1},\cdots,f_dE_d],\\ Q_j^{(1)} &= q[T_c, T_{c+1},\cdots,T_d], & & (j = 1,2,\cdots,r), \end{aligned}$$

where $c = s_1 + s_2 + \cdots s_{j-1} + 1$, $d = s_1 + s_2 + \cdots + s_j$. Using this notation we may express our last result in the form of a lemma;

LEMMA I. *If the two canonical forms QG and QF of equations* (37) *and* (38) *respectively, are equivalent, there exist $2r$ non-singular matrices $D_j(\theta)$, $D^*_j(\theta)$ with elements in $K(\theta,i)$ such that,*

$$(51)\qquad D^*_j(\theta)\gamma_j(\theta)D_j(\theta) = \phi_j(\theta), \qquad (j = 1,2,\cdots,r);$$

i. e., *the matrices $\gamma_j(\theta)$, $\phi_j(\theta)$, $(j = 1,2,\cdots,r)$, are equivalent under a non-singular conjunctive transformation in the field $K(\theta,i)$.*

The converse of this lemma is also true. In fact (51) implies that

[13] If i lies in the field $K(\theta)$, $|D(\theta)| \neq 0$ does not imply $|D^*(\theta)| \neq 0$.

$\tilde{D}_j\gamma_j D_j = \phi_j$, $(j = 1, 2, \cdots, r)$, where $|D_j| \neq 0$. If W_j is the matrix obtained from D_j by replacing each element d of D_j by $E_{s_1+s_2+\cdots+s_j}$, it immediately follows that

$$\tilde{W}_j \Gamma_j W_j = \Phi_j, \qquad (j = 1, 2, \cdots, r),$$

and since $Q_j^{(1)}\tilde{W}_j = W^*_j Q_j^{(1)}$, that

$$W^*_j Q_j^{(1)} \Gamma_j W_j = Q_j^{(1)} \Phi_j.$$

Hence, if $W = [W_1, W_2, \cdots, W_r]$, W is non-singular and $W^*QGW = QF$, so that the two normal forms QG and QF are equivalent.

In stating the theorem given below we use a notation conforming with that explained in (49) and (50); the matrices defined in (50) are associated with a particular polynomial $p_i(\lambda)$ and for convenience in writing we dropped the suffix i but now we find it necessary to replace it. We have proved the theorem:

THEOREM I. *Let A and B be two matrices, of which the second B is non-singular, with elements in $K(i)$ and let $A = A^*$ and $B = B^*$. If the elementary factors of $A - \lambda B$ are the polynomials $[p_i(\lambda)]^{\eta_{ij}}$ of* (3), *then a canonical form for the pencil $A - \lambda B$ under a non-singular conjunctive transformation is the diagonal block matrix*

$$(52) \quad QG(M - \lambda E) = Q[\Gamma_{ij}(L_{ij} - \lambda I_{ij})], \quad (i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, r_i),$$

where Q is defined by (13) *while L_{ij}, Γ_{ij} and I_{ij} are defined by* (50). *Two canonical forms $QG(M - \lambda E)$ and $QF(M - \lambda E)$, where $F = [\Phi_{ij}]$, $(i = 1, 2, \cdots, t; j = 1, 2, \cdots, r_i)$, are equivalent, if and only if the diagonal matrices $\gamma_{ij}(\theta_i)$ and $\phi_{ij}(\theta_i)$ are equivalent under a conjunctive transformation in the field $K(\theta_i, i)$, $(i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, r_i)$.*

Thus, if $[p_i(\lambda)]^{\xi_{ij}}$ occurs exactly s_{ij} times among the elementary factors of $A - \lambda B$, in a canonical form (52), there is associated with this elementary factor a diagonal matrix $\gamma_{ij}(\theta_i)$ of order s_{ij}, whose elements lie in the field $K(\theta_i)$, where θ_i is a root of the irreducible equation $p_i(\lambda) = 0$. The matrix $\gamma_{ij}(\theta_i)$ is determined apart from a conjunctive transformation.

Throughout we have used conjunctive transformation to include the case of congruent transformation. We therefore see that, *if A and B are symmetric matrices over K, Theorem I is true when 'conjunctive' is replaced by 'congruent' and $K(i)$ is replaced by K.*

We now state two corollaries of Theorem 1.

COROLLARY I. *We may determine the matrices γ_{ij} of a canonical form (52) in such a way that no element of γ_{ij} contains a factor r^2, where r is a polynomial in the matrix p with coefficients in K.*

For, if $\phi_{ij} = \rho_{ij}{}^2\gamma_{ij}$, and ρ_{ij} is a diagonal matrix whose elements are polynomials in p_i with coefficients in K, $\rho^*{}_{ij}\gamma_{ij}\rho_{ij} = \phi_{ij}$, so that ϕ_{ij} is equivalent to γ_{ij}.

COROLLARY II. *Two pairs of hermitian matrices A, B and C, D, with elements in $K(i)$, the second of each pair being non-singular, are equivalent under a non-singular conjunctive transformation in $K(i)$, if, and only if, the two pencils $A - \lambda B$ and $C - \lambda D$ have the same elementary factors and, if the matrices $\gamma_{ij}(\theta_i)$, associated with each distinct elementary factor, are conjunctively equivalent in $K(i, \theta_i)$.*

COROLLARY III. *Corollary II remains true if hermitian is replaced by symmetric, conjunctive by congruent and $K(i)$ by K.*

We may use Theorem 1 to determine a canonical form for any non-singular pencil of matrices $A - \lambda B$ with elements in $K(i)$. For, if B is singular but $| A - \lambda B | \neq 0$, we may determine a new basis for the pencil, A_1 and B_1, where B_1 is non-singular and $A - \lambda B = A_1 - \rho B_1$.[14] We apply Theorem I to the pencil $A_1 - \rho B_1$ and thus determine a canonical form for the non-singular pencil $A - \lambda B$.

3. *Ordinary hermitian matrices and real symmetric matrices.* If K is the field of all real numbers and $K(i)$ the complex number field, the polynomials $p_i(\lambda)$ of (3) are either quadratic or linear. If $p_i(\lambda)$ is quadratic, $K(\theta_i) = K(i)$ and hence, if $\gamma_{ij}(\theta_i)$ is one of the matrices associated with $p_i(\lambda)$, $\gamma_{ij}(\theta_i)$ is a diagonal matrix, whose elements are complex numbers. Let $\gamma_{ij}(\theta_i) = [g_r]$ and let $W = [w_r]$, where $w_r = g_r{}^{1/2}$, if $g_r \neq 0$, and $w_r = 1$, if $g_r = 0$. Then the matrix W is a non-singular matrix with elements in $K(\theta_i)$, as is the matrix W^{-1}. But $(W^{-1})'\gamma_{ij}(\theta_i)W^{-1}$ is the identity matrix. Hence each matrix $\gamma_{ij}(\theta_i)$ associated with $p_i(\lambda)$ may be reduced to the identity matrix of the corresponding order. If $p_i(\lambda) = \lambda^2 - 2a_i\lambda + a_i{}^2 + b_i{}^2$, we may choose for p_i the matrix $\begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}$ and for q_i the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. If, however, $p_i(\lambda)$ is linear, $K(\theta_i) = K$, the field of all real numbers and by Corollary 2 each associated $\gamma_{ij}(\theta_i)$ may be reduced to a diagonal matrix with elements, which are either $+1$ or -1. If $p_i(\lambda) = \lambda - \lambda_i$, $p = \lambda_i$ and

[14] Cf. Turnbull and Aitken, *op. cit.*, p. 117 sq.; Trott 1, p. 370.

$q=1$. The normal form (52) therefore coincides with the normal form given by Trott 1, page 368, formula (11). In this particular case, however, the condition of Lemma 1 is greatly simplified, for two diagonal matrices with real coefficients are equivalent under a conjunctive or congruent transformation, if, and only if, they have the same signature. Thus Trott's condition (15) merely expresses the fact that $\gamma_{ij}(\theta_i)$ is conjunctively equivalent to $\phi_{ij}(\theta_i)$.

In the general case no such simplification of the conditions in Lemma 1 is possible. The conditions for the equivalence of two quadratic or hermitian forms have been determined but are very complicated.[15] We however state necessary and sufficient conditions in the two simplest cases (a) $s_{ij}=1$, (b) $s_{ij}=2$. These conditions are due to Dickson.

(a) If $\gamma_{ij}(\theta_i)$ is of order one, $\gamma_{ij}(\theta_i)$ is an element of $K(\theta_i)$. Then *$\gamma_{ij}(\theta_i)$ is equivalent to $\phi_{ij}(\theta_i)$ if, and only if, there exists an element f of $K(i,\theta_i)$, such that*

$$\phi_{ij}(\theta_i)=f\bar{f}\gamma_{ij}(\theta_i).$$

(b) If the matrices $\gamma_{ij}(\theta_i)$ and $\phi_{ij}(\theta_i)$ are both of order two, they may be represented as $\begin{pmatrix}\gamma_1 & 0\\ 0 & \gamma_2\end{pmatrix}$ and $\begin{pmatrix}\phi_1 & 0\\ 0 & \phi_2\end{pmatrix}$ respectively. Then *$\gamma_{ij}(\theta_i)$ is equivalent to $\phi_{ij}(\theta_i)$, if, and only if, there exist elements f, g and h of $K(\theta,i)$, such that $\phi_1=\gamma_1 f\bar{f}+\gamma_2 g\bar{g}$ and $\phi_1\phi_2=h\bar{h}\gamma_1\gamma_2$.* In the symmetric case the elements f, g, h lie in K and $f=\bar{f}$, $g=\bar{g}$, $h=\bar{h}$.

4. We conclude our discussion by giving an explicit form for the matrices, p_i and q_i, which occur in the canonical form (52). The matrix $p=p_i$ is a matrix of order $n=n_i$, whose characteristic equation is the irreducible equation $p_i(\lambda)=p(\lambda)=0$ of degree n. If

$$p(\lambda)=\lambda^n-a_n\lambda^{n-1}-a_{n-1}\lambda^{n-2}\cdots-a_2\lambda-a_1,$$

we may take for the matrix p the *companion matrix* of $p(\lambda)$,

$$p=\begin{bmatrix}0 & 0 & \cdot & \cdot & \cdot & 0 & a_1\\ 1 & 0 & \cdot & \cdot & \cdot & 0 & a_2\\ 0 & 1 & \cdot & \cdot & \cdot & 0 & a_3\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \cdot & \cdot & \cdot & 1 & a_n\end{bmatrix},$$

[15] L. E. Dickson, "On quadratic, bilinear, and Hermitian forms," *Transactions of the American Mathematical Society*, vol. 7 (1906), pp. 275-292; "On quadratic forms in a general field," *Bulletin of the American Mathematical Society*, vol. 14 (1907-8), pp. 108-115; H. Hasse, "Symmetrische Matrizen in Körper der rationalen zahlen," *Crelle*, vol. 153, pp. 12-43.

and for $q = q_i$ the matrix

$$(53) \qquad q = \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & b_2 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & b_2 & b_3 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & b_2 & b_3 & b_4 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & b_2 & b_3 & \cdot & \cdot & \cdot & b_{n-2} & b_{n-1} & b_n \end{pmatrix},$$

where $b_2 = a_1$ and $b_{i+1} = b_2 a_{n+2-i} + b_3 a_{n+3-i} + \cdots + b_i a_n$, $(i = 2, 3, \cdots, n-1)$. It is obvious that q is non-singular, for p is non-singular and hence $b_2 = a_1 \neq 0$. Moreover it is easily verified that $qp = p'q$. The matrix q is not uniquely determined by the matrix p, but the matrix q in (53) is as simple in form as, if not simpler than, any other matrix q_1 satisfying the equation $p'q_1 = q_1 p$.

THE JOHNS HOPKINS UNIVERSITY.

ON THE RATIONAL CANONICAL FORM OF A FUNCTION OF A MATRIX.

By Neal H. McCoy.

Let A be a matrix of order n with elements in the complex number field, and $\phi(A)$ a given rational integral function of A. In 1906, Kreis [1] gave a method of determining the elementary divisors of $\phi(A)$ from those of A. In recent years the same problem has been discussed by Krishnamurthy,[2] Turnbull and Aitken,[3] Rutherford [4] and Amante.[5] It is sufficient to consider the case in which A has a single elementary divisor $(\lambda - a)^n$, as the general case easily reduces to this one. The principal result of these writers may then be stated in the following way. Expand $\phi(\lambda)$ in powers of $\lambda - a$,

$$\phi(\lambda) = a_0 + a_1(\lambda - a) + a_2(\lambda - a)^2 + \cdots.$$

Suppose the i-th number of the sequence $a_1, a_2, \cdots, a_{n-1}, 1$, is the first which is not zero. Define positive integers k and l by the relations,

$$n = (k-1)i + l, \qquad k \geqq 1, \qquad 1 \leqq l \leqq i.$$

Then $\phi(A)$ has the elementary divisors $(\lambda - a_0)^k$ taken l times, and $(\lambda - a_0)^{k-1}$ taken $i - l$ times.

So far as the writer is aware, no solution has been given of the problem corresponding to this one, for the case in which the elements of A and all operations are restricted to an arbitrary domain of rationality. In this more general problem one does not have the use of the comparatively simple Jordan normal form of a matrix, and a different method of attack must therefore be used. It is the purpose of the present paper to present a solution of this problem.

In § 4, we shall also give a brief account of an application of the main result to the solution of certain matric equations.

1. The rational canonical form. Let K denote a given field. Unless

[1] H. Kreis, *Contribution à la théorie des systèmes linéaires*, Zürich, 1906.

[2] Rao S. Krishnamurthy, "Invariant-factors of a certain class of linear substitutions," *Journal of the Indian Mathematical Society*, vol. 19 (1932), pp. 233-240.

[3] H. W. Turnbull and A. C. Aitken, *Canonical Matrices*, Glasgow, 1932, pp. 75-76.

[4] D. E. Rutherford, "On the canonical form of a rational integral function of a matrix," *Proceedings of the Edinburgh Mathematical Society* II, vol. 3 (1932), pp. 135-143.

[5] S. Amante, "Sulle riduzione a forma canonica di una classe speciale di matrici," *Atti della Reale Accademia Nazionale dei Lincei*, Rendiconti VI, vol. 17 (1933), pp. 31-36 and pp. 431-436.

otherwise stated, it will be assumed henceforth that all matrices and vectors have coördinates in K, and all polynomials have coefficients in K. If a polynomial is irreducible relative to the field K, we shall simply say that it is irreducible.

Let $f(\lambda) = \lambda^p - a_1\lambda^{p-1} - \cdots - a_p$ be a given polynomial, and form the matrix,

$$B = \begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_p & a_{p-1} & a_{p-2} & \cdot & \cdot & \cdot & a_1 \end{bmatrix}.$$

This matrix may be called the *companion matrix* of the function $f(\lambda)$ or of the equation $f(\lambda) = 0$.[6] The minimum function of B is then $|\lambda - B| = f(\lambda)$.

Let now A be a given matrix of order n, and $E_i(\lambda)$, $(i = 1, 2, \cdots, r)$, the non-constant invariant factors of $\lambda - A$. Perhaps the most common rational canonical form of A (with respect to similarity transformations) is a matrix A_1, which is the direct sum[7] of the companion matrices of the $E_i(\lambda)$. However, it will be convenient for our purpose to use a somewhat different rational canonical form, which will now be described.

If we factor the $E_i(\lambda)$ into powers of distinct, irreducible polynomials $p_k(\lambda)$, each of which has leading coefficient unity, say

$$E_i(\lambda) = [p_1(\lambda)]^{n_{i1}}[p_2(\lambda)]^{n_{i2}} \cdots [p_t(\lambda)]^{n_{it}} \qquad (i = 1, 2, \cdots, r),$$

then such of the factors $[p_k(\lambda)]^{n_{ij}}$ as are not mere constants may be called the *elementary divisors* of A. We can then choose as a canonical form of A, a matrix A_2, which is the direct sum of the companion matrices of the elementary divisors of A.[8] This is the canonical form used throughout this paper. The advantage of this form over the other lies in the fact that if $A = C \dotplus D$, the canonical form of A is the direct sum of the canonical forms of C and D, and the elementary divisors of A are the elementary divisors of C, together with those of D.

---

[6] See C. C. MacDuffee, *The Theory of Matrices*, Berlin, 1933, p. 20.

[7] If $A = \begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$, where C and D are square matrices, then A is called the *direct sum* of C and D, and we write, $A = C \dotplus D$.

[8] W. Krull, "Theorie und Anwendung der verallgemeinerten Abelschen Gruppen," *Sitzungsberichte Heidelberger Akademie der Wissenschaften*, 1926, pp 25-28; B. L. van der Waerden, *Moderne Algebra*, vol. 2, Berlin, 1931, p. 137. For a somewhat different canonical form, but one which also uses the notion of elementary divisors, see J. H. M. Wedderburn, "Note on matrices in a given field," *Annals of Mathematics*, vol. 27 (1926), pp. 245-248.

We shall now establish two lemmas [9] which will be useful also in a later section of the paper, and then apply them to show how to find a non-singular matrix which transforms A_1 into A_2.

LEMMA 1. *Let V_j be a given row vector of dimension n, X an arbitrary column vector of dimension n, and B a given square matrix of order n. Denote by R_j the matrix of e_j rows and n columns, whose rows are respectively the vectors*

$$V_j, V_jB, V_jB^2, \cdots, V_jB^{e_j-1}.$$

If now $h_j(\lambda) = \lambda^{e_j} - b_1\lambda^{e_j-1} - \cdots - b_{e_j}$, is a polynomial such that $V_jh_j(B) = 0$, and we set $\xi_j = R_jX$, $Y = BX$, $\eta_j = R_jY$, then it follows that $\eta_j = Q_j\xi_j$, where Q_j is the companion matrix of $h_j(\lambda)$.

By definition, we see that $\xi_j = \{\xi_{j1}, \xi_{j2}, \cdots, \xi_{je_j}\}$ and $\eta_j = \{\eta_{j1}, \eta_{j2}, \cdots, \eta_{je_j}\}$ are column vectors of dimension e_j. The lemma follows at once from the following calculation:

$$\begin{aligned}
V_jX &= \xi_{j1},\\
\eta_{j1} = V_jBX &= \xi_{j2},\\
\eta_{j2} = V_jB^2X &= \xi_{j3},\\
\cdots & \cdots\\
\eta_{j,e_j-1} = V_jB^{e_j-1}X &= \xi_{j,e_j}\\
\eta_{j,e_j} = V_jB^{e_j}X &= b_1\xi_{j,e_j-1} + b_2\xi_{j,e_j-2} + \cdots + b_{e_j}\xi_{j1}.
\end{aligned}$$

LEMMA 2. *Let e_j $(j = 1, 2, \cdots, q)$ be positive integers whose sum is n, and for each j suppose V_j, R_j, ξ_j, η_j, $h_j(\lambda)$, Q_j, satisfy the conditions of Lemma 1. If we set*

$$R = \begin{pmatrix} R_1 \\ R_2 \\ \cdot \\ R_q \end{pmatrix}, \qquad \xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \xi_q \end{pmatrix}, \qquad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \cdot \\ \eta_q \end{pmatrix},$$

then $\xi = RX$, $Y = BX$, $\eta = RY$, and $\eta = Q\xi$, where $Q = Q_1 \dot{+} Q_2 \dot{+} \cdots \dot{+} Q_q$. Further, if R is non-singular, then $Q = RBR^{-1}$.

The first part follows almost immediately from the preceding lemma. We then find that $RBX = QRX$. But since X is entirely arbitrary, we must have $RB = QR$. Hence if R is non-singular, $Q = RBR^{-1}$.

---

[9] I am indebted to a referee for suggesting the introduction of these lemmas. Their use has considerably improved the proof of Theorem 1.

Let $U = (u_1, u_2, \cdots, u_n)$ be any row vector. Then there exists a unique polynomial $g(\lambda)$ of minimum degree and with leading coefficient unity, such that $Ug(A) = 0$.[10] This polynomial $g(\lambda)$ is called the *R.C.F.* (*Reduced Characteristic Function*) of A relative to U, and its degree may be called the *grade* of U (relative to A). A fundamental property of the R. C. F. of A relative to U is that, if $h(\lambda)$ is a polynomial such that $Uh(A) = 0$, then $h(\lambda)$ is divisible by $g(\lambda)$.

We now return to the problem of transforming the matrix A_1 into A_2. Since A_1 is the direct sum of the companion matrices of the invariant factors of $\lambda - A$, we may assume for our purpose, that $\lambda - A$ has a single invariant factor $E(\lambda)$ of degree n. If $E(\lambda)$ is a power of an irreducible polynomial, then clearly $A_1 = A_2$. Hence suppose $E(\lambda) = \phi(\lambda)\psi(\lambda)$, where $\phi(\lambda)$ and $\psi(\lambda)$ are relatively prime, and have leading coefficients equal to unity. Let the degrees of $\phi(\lambda)$ and $\psi(\lambda)$ be respectively n_1 and n_2.

From the form of A_1, it follows that the vector, $U_1 = (1, 0, \cdots, 0)$, is of grade n relative to A_1, and the R. C. F. of A_1 relative to U_1 is therefore $E(\lambda)$. We may now apply the above lemmas by placing $e_1 = n_2$, $e_2 = n_1$, $B = A_1$, $V_1 = U_1\phi(A_1)$, $V_2 = U_1\psi(A_1)$. Since $U\phi(A_1)\psi(A_1) = 0$, it follows at once that $h_1(\lambda) = \psi(\lambda)$, $h_2(\lambda) = \phi(\lambda)$. If now we assume for the moment that R is non-singular, we see by means of Lemma 2 that $RA_1R^{-1} = Q$, where Q is the direct sum of the companion matrices of $\phi(\lambda)$ and of $\psi(\lambda)$. If either $\phi(\lambda)$ or $\psi(\lambda)$ can be expressed as a product of relatively prime factors, the process can be continued, and so on until the form A_2 is reached.

We now show that R is non-singular. For suppose there exists a relation

$$U_1[\sum_{i=0}^{n_2-1} c_i\phi(A_1)A_1^i + \sum_{i=0}^{n_1-1} d_i\psi(A_1)A_1^i] = 0.$$

Since $E(\lambda)$ is the R. C. F. of U_1 relative to A_1, this implies that the polynomial

$$F(\lambda) \equiv \phi(\lambda)\sum_{i=0}^{n_2-1} c_i\lambda^i + \psi(\lambda)\sum_{i=0}^{n_1-1} d_i\lambda^i,$$

is divisible by $E(\lambda)$, and being of degree at most $n-1$, must therefore vanish identically. From the fact that $\phi(\lambda)$ and $\psi(\lambda)$ are relatively prime, it follows that all coefficients c_i and d_i must be zero, and thus R is non-singular.

2. Another lemma. Let $p(\lambda)$ and $\phi(\lambda)$ be given polynomials, of which the first is irreducible and of degree $s \geqq 1$. Since there exist at most s polynomials which are linearly independent modulo $p(\lambda)$, it follows that the polynomials $1, \phi, \phi^2, \cdots, \phi^s$ are linearly dependent modulo $p(\lambda)$. By a

[10] Turnbull and Aitken, *op. cit.*, chap. 6.

familiar argument, there then exists a unique polynomial $f(x)$, with leading coefficient unity, and of minimum degree, such that

$$f(\phi(\lambda)) \equiv 0 \qquad (\operatorname{mod} p(\lambda)).$$

It follows readily that $f(x)$ is irreducible, and also that if $g(x)$ is a polynomial, such that $g(\phi(\lambda)) \equiv 0 \ (\operatorname{mod} p(\lambda))$, then $g(x) \equiv 0 \ (\operatorname{mod} f(x))$. We shall let t denote the degree of $f(x)$.

Let ρ denote a root of $p(\lambda) = 0$ in a properly extended field, and consider the three fields,

$$K \subseteqq K(\phi(\rho)) \subseteqq K(\rho).$$

The field $K(\phi(\rho))$ is seen to be of degree t over K, as $\phi(\rho)$ satisfies the irreducible equation $f(x) = 0$. Also $K(\rho)$ is of degree s over K. Hence, by a well known theorem,[11] $K(\rho)$ is algebraic of degree $m = s/t$ over the field $K(\phi(\rho))$. That is, t is a divisor of s, and ρ satisfies no equation of degree less than m with coefficients in $K(\phi(\rho))$. We may now prove the following lemma:

LEMMA 3. *If $F(x, y)$ is a polynomial in the indeterminates x, y, of degree at most $m - 1$ in y, and if*

$$F(\phi(\lambda), \lambda) \equiv 0 \qquad (\operatorname{mod} p(\lambda)),$$

then

$$F(x, y) \equiv 0 \qquad (\operatorname{mod} f(x)).$$

Under the hypotheses of the lemma, we have $F(\phi(\rho), \rho) = 0$. Let $F(x, y) = \sum_{i=0}^{m-1} F_i(x) y^i$. We have then $\sum_{i=0}^{m-1} F_i(\phi(\rho))\rho^i = 0$. But if some $F_i(\phi(\rho)) \neq 0$, this contradicts the fact that ρ can satisfy no equation of degree less than m with coefficients in $K(\phi(\rho))$. Hence $F_i(\phi(\rho)) = 0$, and thus $F_i(\phi(\lambda)) \equiv 0 \ (\operatorname{mod} p(\lambda))$, $(i = 0, 1, \cdots, m - 1)$. It follows that each $F_i(x) \equiv 0 \ (\operatorname{mod} f(x))$, and the lemma is established. We remark that if the degree of $F(x, y)$ in x is at most $t - 1$, then $F(x, y)$ vanishes identically.

3. The elementary divisors of $\phi(A)$. We come now to the main problem of the paper. Let A be a given matrix, and $\phi(A)$ a given polynomial in A. Since $\phi(HAH^{-1}) = H\phi(A)H^{-1}$, there is no loss of generality in assuming that A is in canonical form. If $A = A_1 \dotplus A_2$, then $\phi(A) = \phi(A_1) \dotplus \phi(A_2)$, and the elementary divisors of $\phi(A)$ are precisely those of $\phi(A_1)$, together with those of $\phi(A_2)$. We shall therefore assume henceforth that A has a single elementary divisor $[p(\lambda)]^r$. It follows that the minimum function of A is $[p(\lambda)]^r$. If the degree of $p(\lambda)$ is s, then the order of A is $n = rs$.

[11] See, e. g., van der Waerden, *op. cit.*, vol. 1, p. 98.

Let $f(x)$ denote the unique irreducible polynomial of degree t defined in the preceding section. Then we have

$$f(\phi(\lambda)) \equiv 0 \qquad (\operatorname{mod} p(\lambda)).$$

It may well happen that $f(\phi(\lambda))$ is divisible by a power of $p(\lambda)$ greater than the first. Suppose that it is divisible by $[p(\lambda)]^q$ but not by $[p(\lambda)]^{q+1}$. We now define an integer i as follows. If $q \geqq r$, we set $i = r$; while if $q < r$, we place $i = q$. Hence in either case we have $f(\phi(\lambda)) \equiv 0 \pmod{[p(\lambda)]^i}$. We further define positive integers k, l by the relations,

$$(1) \qquad r = (k-1)i + l, \qquad k \geqq 1, \qquad 1 \leqq l \leqq i.$$

It follows that $[f(\phi(\lambda))]^k \equiv 0 \pmod{[p(\lambda)]^r}$, while

$$[f(\phi(\lambda))]^{k-1} \not\equiv 0 \qquad (\operatorname{mod} [p(\lambda)]^r).$$

The minimum function of $\phi(A)$ is therefore $[f(\lambda)]^k$, and the elementary divisors of $\phi(A)$ are all powers of $f(\lambda)$. If we denote the integer s/t by m, we may state the following precise result:

THEOREM 1. *The matrix $\phi(A)$ has as elementary divisors, $[f(\lambda)]^k$ taken lm times, and $[f(\lambda)]^{k-1}$ taken $m(i-l)$ times.*

We shall prove this theorem by actually exhibiting a matrix R which transforms $\phi(A)$ to canonical form. Let U denote a vector of grade $n = rs$ with respect to A.[12] The R. C. F. of A relative to U is then $[p(\lambda)]^r$.

Let α and β be integers such that $0 \leqq \alpha \leqq i-1$, $0 \leqq \beta \leqq m-1$. We shall now make use of Lemma 1, the notation being as in the statement of the lemma, with the exception that we shall find it convenient to replace each subscript j by the two subscripts α and β. That is, e_j becomes $e_{\alpha\beta}$, $h_j(\lambda)$ becomes $h_{\alpha\beta}(\lambda)$, and so on. Two cases will be considered separately.

Case 1. $0 \leqq \alpha \leqq l-1$, $0 \leqq \beta \leqq m-1$. Let $e_{\alpha\beta} = tk$, $V_{\alpha\beta} = U[p(A)]^\alpha A^\beta$, $B = \phi(A)$. Since $U[p(A)]^\alpha A^\beta [f(\phi(A))]^k = 0$, it follows that $h_{\alpha\beta}(\lambda) = [f(\lambda)]^k$, and $Q_{\alpha\beta}$ is therefore the companion matrix of $[f(\lambda)]^k$. The matrix $R_{\alpha\beta}$ has as rows the vectors,

$$(2) \quad U[p(A)]^\alpha A^\beta, \quad U[p(A)]^\alpha A^\beta \phi(A), \cdots, U[p(A)]^\alpha A^\beta [\phi(A)]^{tk-1}.$$

Case 2. $l \leqq \alpha \leqq i-1$, $0 \leqq \beta \leqq m-1$. In this case, let $e_{\alpha\beta} = t(k-1)$, $V_{\alpha\beta} = U[p(A)]^\alpha A^\beta$, $B = \phi(A)$. Since now $f(\phi(\lambda))$ is divisible by $[p(\lambda)]^i$,

[12] If A is in canonical form, we may choose $U = (1, 0, \cdots, 0)$, as in § 1.

it follows by relations (1), that $U[p(A)]^{\alpha}A^{\beta}[f(\phi(A))]^{k-1}=0$, and thus that $h_{\alpha\beta}(\lambda)=[f(\lambda)]^{k-1}$. The matrix $R_{\alpha\beta}$ has as rows the vectors,

$$(3)\quad U[p(A)]^{\alpha}A^{\beta},\ U[p(A)]^{\alpha}A^{\beta}\phi(A),\cdots,U[p(A)]^{\alpha}A^{\beta}[\phi(A)]^{t(k-1)-1}.$$

It is easily seen that $\sum_{\alpha=0}^{i-1}\sum_{\beta=0}^{m-1}e_{\alpha\beta}=n$, and the hypotheses of Lemma 2 are satisfied. The matrix R then is formed by arranging the matrices $R_{\alpha\beta}$ ($\alpha=0,1,\cdots,i-1$; $\beta=0,1,\cdots,m-1$) in some fixed order, and using the same order for the $\xi_{\alpha\beta}$ and $\eta_{\alpha\beta}$ to define ξ and η as in the statement of the lemma. Let us now assume for the present that R is non-singular. We then have $Q=R\phi(A)R^{-1}$, and by the determinations of $h_{\alpha\beta}(\lambda)$ above, we see that Q is the direct sum of the companion matrix of $[f(\lambda)]^{k}$ taken ml times, and of the companion matrix of $[f(\lambda)]^{k-1}$ taken $m(i-l)$ times. But since $f(\lambda)$ is irreducible, Q is therefore the canonical form of $\phi(A)$, and the elementary divisors are those stated in the theorem.

There remains only to prove that R is non-singular. Any linear combination $UF(\phi(A),A)$ of the row vectors of R (of the types (2) and (3)) corresponds to a polynomial $F(x,y)$ of the form

$$(4)\qquad F(x,y)=\sum_{j=0}^{i-1}F_j(x,y)\,[p(y)]^{j},$$

where the degree of $F_j(x,y)$ is at most $m-1$ in y, while its degree in x is at most $tk-1$ for $j=0,1,\cdots,l-1$, and at most $t(k-1)-1$, for $j=l,l+1,\cdots,i-1$. If the linear combination of the rows of R is the zero vector, we have

$$UF(\phi(A),A)=0,$$

and since the R. C. F. of A relative to U is $[p(\lambda)]^{r}$, it follows that

$$(5)\qquad F(\phi(\lambda),\lambda)\equiv 0\qquad (\mathrm{mod}\ [p(\lambda)]^{r}).$$

We shall complete the proof by showing that under these conditions, all $F_j(x,y)$ vanish identically, and thus the rows of R are linearly independent.

We first dispose of the special case in which $i=r$, and hence $k=1$, $l=r$. In this case, all $F_j(x,y)$ are of degree at most $t-1$ in x. From relation (5), we find that

$$\sum_{j=0}^{r-1}F_j(\phi(\lambda),\lambda)[p(\lambda)]^{j}\equiv 0\qquad (\mathrm{mod}\ [p(\lambda)]^{r}).$$

Now clearly $F_0(\phi(\lambda),\lambda)\equiv 0$ (mod $p(\lambda)$), and by Lemma 3, it follows that $F_0(x,y)\equiv 0$ (mod $f(x)$). But being of degree at most $t-1$ in x, $F_0(x,y)$

must vanish identically. We now pass on to $F_1(x, y)$, and a similar argument shows that it is also identically zero. A continuation of this process establishes the fact that all $F_j(x, y)$ vanish identically.

Suppose now that $i < r$. By definition of i, we know that $f(\phi(\lambda))$ is then divisible by $[p(\lambda)]^i$ but not by $[p(\lambda)]^{i+1}$. We now assume that all $F_j(x, y)$ are divisible by $[f(x)]^\gamma$ where $0 \leqq \gamma < k-1$, and shall show that they are all divisible by $[f(x)]^{\gamma+1}$. If we set $F_j(x, y) = [f(x)]^\gamma F'_j(x, y)$, we get from (5),

$$(6) \qquad \sum_{j=0}^{i-1} F'_j(\phi(\lambda), \lambda)[p(\lambda)]^j \equiv 0 \qquad (\text{mod } [p(\lambda)]^{r-\gamma i}).$$

Clearly $F'_0(\phi(\lambda), \lambda) \equiv 0 \pmod{p(\lambda)}$, and by Lemma 3 we have $F'_0(x, y) \equiv 0 \pmod{f(x)}$. Suppose that $F'_j(x, y) \equiv 0 \pmod{f(x)}$, $(j = 0, 1, \cdots, \delta)$ where $0 \leqq \delta < i-1$. Since $r - \gamma i > i$, it follows that $F'_{\delta+1}(\phi(\lambda), \lambda) \equiv 0 \pmod{p(\lambda)}$, and thus $F'_{\delta+1}(x, y) \equiv 0 \pmod{f(x)}$. Hence $F'_j(x, y) \equiv 0 \pmod{f(x)}$, $(j = 0, 1, \cdots, i-1)$. It therefore follows that all $F_j(x, y)$ are divisible by $[f(x)]^{\gamma+1}$, and a process of induction then shows that they are all divisible by $[f(x)]^{k-1}$. But the $F_j(x, y)$ $(j = l, l+1, \cdots, i-1)$ are of degree at most $t(k-1) - 1$ in x, and hence must vanish identically.

Now let $F_j(x, y) = [f(x)]^{k-1} F^*_j(x, y)$, $(j = 0, 1, \cdots, l-1)$. From relation (5) we then have

$$(7) \qquad \sum_{j=0}^{l-1} F^*_j(\phi(\lambda), \lambda)[p(\lambda)]^j \equiv 0 \qquad (\text{mod } [p(\lambda)]^l).$$

A repetition of the argument of the preceding paragraphs shows that each $F^*_j(x, y)$ is divisible by $f(x)$, and thus $F_j(x, y)$ is divisible by $[f(x)]^k$, $(j = 0, 1, \cdots, l-1)$. But these $F_j(x, y)$ are of degree at most $tk - 1$ in x, and must therefore vanish identically. This completes the proof of the theorem.

Examples. Let K be an algebraically closed field, and suppose A has the single elementary divisor $(\lambda - a)^n$. Then in terms of our notation, we have $p(\lambda) = \lambda - a$, $s = 1$, $r = n$. Let now $\phi(\lambda)$ be expanded in powers of $\lambda - a$,

$$\phi(\lambda) = a_0 + a_1(\lambda - a) + a_2(\lambda - a)^2 + \cdots.$$

Then clearly $\phi(\lambda) - a_0 \equiv 0 \pmod{(\lambda - a)}$, so that $f(x) = x - a_0$, and $t = 1$, $m = s = 1$. If now the first number of the sequence $a_1, a_2, \cdots, a_{n-1}, 1$, which is not zero is the i-th, then we have

$$f(\phi(\lambda)) \equiv 0 \qquad (\text{mod } (\lambda - a)^i),$$

while if $i < n$,

$$f(\phi(\lambda)) \not\equiv 0 \qquad (\text{mod } (\lambda - a)^{i+1}).$$

Thus this definition of i corresponds to that given in the notation above, and if we define k and l by the relations (1), our theorem tells us that $\phi(A)$ has the elementary divisor $(\lambda - a_0)^k$ taken l times and $(\lambda - a_0)^{k-1}$ taken $(i - l)$ times. Thus our general theorem reduces to the one obtained previously for this case by the writers referred to in the introduction.

As a second example, let K be the field of real numbers, and A a matrix of order $n = 6$ with the elementary divisor $(\lambda^2 + 1)^3$. We may take A in the canonical form,

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & -3 & 0 & -3 & 0 \end{pmatrix}.$$

Then $p(\lambda) = \lambda^2 + 1$, $s = 2$, $r = 3$. Suppose $\phi(\lambda) = \lambda^3 + 3\lambda$. Then $\phi(\lambda) \equiv 2\lambda \pmod{p(\lambda)}$, $[\phi(\lambda)]^2 \equiv -4 \pmod{p(\lambda)}$, and hence $f(x) = x^2 + 4$. We have then $t = 2$, $m = 1$. It is easily verified that $\phi^2 + 4 \equiv 0 \pmod{[p(\lambda)]^2}$, but $\not\equiv 0 \pmod{[p(\lambda)]^3}$. Hence $i = 2$, $k = 2$, $l = 1$. Our theorem then states that $\phi(A) = A^3 + 3A$ has the elementary divisors $(\lambda^2 + 4)^2$, $\lambda^2 + 4$. Thus the canonical form of $\phi(A)$ is the matrix

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -16 & 0 & -8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -4 & 0 \end{pmatrix}.$$

The vector $U = (1, 0, 0, 0, 0, 0)$ is of grade 6 relative to A, and so the matrix R is a matrix whose rows are respectively U, $U\phi(A)$, $U[\phi(A)]^2$, $U[\phi(A)]^3$, $U(A^2 + 1)$, $U(A^2 + 1)\phi(A)$. A calculation shows that

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 & 0 \\ -1 & 0 & 6 & 0 & 3 & 0 \\ 0 & -6 & 0 & 8 & 0 & 6 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 4 & 0 & 1 \end{pmatrix}.$$

It is easily verified that $R\phi(A) = QR$, and hence that $Q = R\phi(A)R^{-1}$.

4. Solution of matric equations. Let B be a given matrix of order n, and $\phi(\lambda)$ a given polynomial in the scalar variable λ. It will be understood that all elements and operations are to be restricted to the given field K. We shall now give a brief account of an application of the results of the preceding section to the solution of the equation,

$$\phi(X) = B, \tag{8}$$

where X is a matrix of order n to be determined. A different method of solving this equation has recently been given by Ingraham.[13]

If S is a non-singualr matrix, then $\phi(SXS^{-1}) = S\phi(X)S^{-1} = SBS^{-1}$. Hence there is no loss of generality in assuming that B is in canonical form. We shall assume henceforth that B is in canonical form, and is therefore the direct sum of the companion matrices of its elementary divisors. We observe that, if X is a solution of the equation (8), then SXS^{-1} is also a solution, if and only if S is commutative with B.

We shall consider first the case in which the elementary divisors of B are all powers of a single irreducible polynomial $f(\lambda)$. Let

$$f(\phi(\lambda)) = a[p_1(\lambda)]^{n_1}[p_2(\lambda)]^{n_2} \cdots [p_k(\lambda)]^{n_k}, \tag{9}$$

be the decomposition of $f(\phi(\lambda))$ into powers of its distinct irreducible factors, each with leading coefficient unity. It follows easily that, if $p(\lambda)$ is any irreducible polynomial, with leading coefficient unity, then $f(x)$ is the unique minimum polynomial (defined in § 2) such that

$$f(\phi(\lambda)) \equiv 0 \qquad (\mathrm{mod}\, p(\lambda)),$$

if and only if $p(\lambda)$ is one of the $p_i(\lambda)$ occurring in (9).

Let X denote a solution of the equation (8), and Y the canonical form of X. That is, $Y = Y_1 \dotplus Y_2 \dotplus \cdots \dotplus Y_r$, where the Y_i are the companion matrices of the elementary divisors of X. Then

$$\phi(Y) = \phi(Y_1) \dotplus \phi(Y_2) \dotplus \cdots \dotplus \phi(Y_r)$$

is similar to B, and the elementary divisors of B are precisely the elementary divisors of all the $\phi(Y_i)$. But by the results of the preceding section, $\phi(Y_i)$ can have elementary divisors which are powers of $f(\lambda)$, if and only if Y_i is the companion matrix of some power of a $p_i(\lambda)$ occurring in (9). Suppose then that the elementary divisors of Y are

[13] M. H. Ingraham, "On the rational solutions of the matrix equation $P(X) = A$," *Journal of Mathematics and Physics*, vol. 13 (1934), pp. 46-50. For additional references to matric equations see MacDuffee, *op. cit.*, chap. 8.

$$
(10) \qquad \begin{array}{l}
[p_1(\lambda)]^{n_{11}}, \cdots, [p_1(\lambda)]^{n_{1t_1}}, \\
[p_2(\lambda)]^{n_{21}}, \cdots, [p_2(\lambda)]^{n_{2t_2}}, \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
[p_k(\lambda)]^{n_{k1}}, \cdots, [p_k(\lambda)]^{n_{kt_k}},
\end{array}
$$

where $n_{ij} \geqq n_{il}$ if $l > j$. Let the degree of $p_i(\lambda)$ be denoted by N_i. Then we must have

$$
(11) \qquad (n_{11} + n_{12} + \cdots + n_{1t_1})N_1 + \cdots + (n_{k1} + n_{k2} + \cdots + n_{kt_k})N_k = n.
$$

We are now in a position to give a method of finding all solutions of the equation (8). Form the diophantine equation (11), and solve it for the n_{ij} under the condition that $n_{ij} \geqq n_{il}$ if $l > j$. Each solution gives us the elementary divisors (10) of a matrix, which is a possible solution of our equation. Form the matrix Y, which is the direct sum of the companion matrices of these elementary divisors. Then by Theorem 1, it is easy to find the elementary divisors of $\phi(Y)$. If these elementary divisors are not the same as the elementary divisors of B, this Y is discarded. If, however, the elementary divisors are identical, then $\phi(Y)$ is similar to B, and the proof of Theorem 1 shows how to find a matrix R such that $R\phi(Y)R^{-1} = B$. If we let $X = RYR^{-1}$, then X is a solution of the equation (8). If $X_1, X_2, \cdots, X_q$ is a complete set of dissimilar solutions, all of which can be found by this method, then the most general solutions are of the form LX_iL^{-1}, where L is a non-singular matrix commutative with B.

It is not difficult to write out additional equations, which together with the equation (11) will serve to determine completely the admissible matrices Y, but the tentative procedure outlined above is perhaps as easy to apply in any given case.

We now return to the general case in which the elementary divisors of B are unrestricted. Suppose these elementary divisors are

$$
(12) \qquad \begin{array}{l}
[f_1(\lambda)]^{m_{11}}, \cdots, [f_1(\lambda)]^{m_{1r_1}}, \\
[f_2(\lambda)]^{m_{21}}, \cdots, [f_2(\lambda)]^{m_{2r_2}}, \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
[f_l(\lambda)]^{m_{l1}}, \cdots, [f_l(\lambda)]^{m_{lr_l}},
\end{array}
$$

where the $f_i(\lambda)$ $(i = 1, 2, \cdots, l)$ are distinct and irreducible. We may then write $B = B_1 \dotplus B_2 \dotplus \cdots \dotplus B_l$, where B_i is the direct sum of the companion matrices of the elementary divisors occurring in the i-th row of the table (12). We shall now prove the following theorem:

THEOREM 2. *If X is any solution of the equation* (8), *then*

$$X = X_1 \dot{+} X_2 \dot{+} \cdots \dot{+} X_l,$$

where X_i is of the same order as B_i, and is a solution of the equation, $\phi(X) = B_i$ $(i = 1, 2, \cdots, l)$.

Let

$$(13) \qquad f_i(\phi(\lambda)) = b_i[p_{i1}(\lambda)]^{a_{i1}}[p_{i2}(\lambda)]^{a_{i2}} \cdots [p_{is_i}(\lambda)]^{a_{is_i}}, \qquad (i = 1, 2, \cdots, l)$$

be the decomposition of the $f_i(\phi(\lambda))$ into powers of distinct irreducible factors, each with leading coefficient unity. If X is a given solution of equation (8), it follows by an argument similar to that used above that the elementary divisors of X are all powers of the $p_{ij}(\lambda)$ $(i = 1, 2, \cdots, l;\ j = 1, 2, \cdots, s_i)$. Let Y_i denote the direct sum of the companion matrices of the elementary divisors of X which are powers of the functions $p_{i1}(\lambda), \cdots, p_{is_i}(\lambda)$ $(i = 1, 2, \cdots, l)$, and set $Y = Y_1 \dot{+} Y_2 \dot{+} \cdots \dot{+} Y_l$. Since the $f_i(\lambda)$ are distinct, it follows that the $p_{ij}(\lambda)$ are all distinct, and the elementary divisors of $\phi(Y)$ which are powers of $f_i(\lambda)$ are precisely the elementary divisors of $\phi(Y_i)$. Hence Y_i is of the same order as B_i, and $\phi(Y_i)$ is similar to B_i $(i = 1, 2, \cdots, l)$. Let us set $SXS^{-1} = Y$, $T_i\phi(Y_i)T_i^{-1} = B_i$, $T = T_1 \dot{+} T_2 \dot{+} \cdots + T_l$. We have then $T\phi(Y)T^{-1} = B$, from which it follows that $\phi(TSXS^{-1}T^{-1}) = B$, and thus TS is commutative with B. It is then known [14] that TS is of the form $M_1 \dot{+} M_2 \dot{+} \cdots \dot{+} M_l$, where M_i is of the same order as B_i, and is commutative with B_i. A calculation shows that

$$X = M_1^{-1}T_1Y_1T_1^{-1}M_1 \dot{+} \cdots \dot{+} M_l^{-1}T_lY_lT_l^{-1}M_l.$$

We find also that $\phi(M_i^{-1}T_iY_iT_i^{-1}M_i) = M_i^{-1}T_i\phi(Y_i)T_i^{-1}M_i = M_i^{-1}B_iM_i = B_i$. The theorem is therefore established.

By means of this theorem, the solution of the general equation (8) is seen to reduce to the solution of a set of equations of the comparatively simple type, in which the elementary divisors of B are all powers of a single irreducible polynomial. Thus all solutions can be found by the method discussed earlier in this section.

SMITH COLLEGE,
NORTHAMPTON, MASS.

[14] See O. Schreier and B. L. van der Waerden, "Die Automorphismen der projektiven Gruppen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 6 (1928), p. 308.

ON CERTAIN TYPES OF HEXAGONS.[1]

By J. R. Musselman.

1. The resolvent, $V_1 = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots + \epsilon^{n-1} x_{n-1}$, where ϵ is a primitive n-th root of unity, was introduced by Lagrange[2] in his memoirs devoted to the fundamental principles of the solutions of the cubic and quartic equations. Its entrance, however, into the field of geometry is very recent. If we represent any point P in the plane by the single complex number p, and if M_k $(k = 0, 1, \cdots, n-1)$ represents the n vertices of a positively-ordered polygon, then when the coördinates p_k of these vertices are subject to one and only one condition, namely that

$$V_1 = \sum_{k=0}^{n-1} \epsilon^k P_k = 0$$

we obtain a polygon which we shall call a positive n-gon of type M. R. L. Echols[3] has used these polygons in giving geometric pictures of the solutions of the cubic and quartic equations. The writer[4] has pointed out recently two different constructions in which these n-gons of type M occur.

In addition to the above studies of this particular type of n-gon, the Lagrange resolvents (and their conjugates) have been used by L. M. Blumenthal[5] to prove that the norm-area of a $2n$-gon is unaltered by translating either of its component n-gons. The Morleys[6] in their recent book have shown that under homologies the $n-1$ Lagrange resolvents for a n-gon form a complete system of relative invariants, and have used them in considering some special ordered n-points. In this article, the Lagrange resolvents are used to disclose some new facts about a well-known figure, to characterize certain interesting ordered six-points, and to prove that connected with any

[1] Read before the National Academy of Science, November 19, 1934.

[2] *Memoirs of Berlin Academy*, 1769; reprinted in *Oeuvres de Lagrange* (Paris, 1868), vol. 3, p. 207.

[3] *The Roots of Circulants and Application to the Roots of Polynomials*. The University of Virginia (1928).

[4] "On certain types of polygons," *The American Mathematical Monthly*, vol. 40 (1933), p. 157.

[5] "Lagrange resolvents in Euclidean geometry," *The American Journal of Mathematics*, vol. 49 (1927), p. 511.

[6] *Inversive Geometry*, G. Bell & Sons, London (1933), p. 203.

six points there are two circumscribed hexagons, whose opposite sides are parallel and whose vertices lie on rectangular hyperbolas.

2. If on the sides of any triangle $A_1A_2A_3$ we construct the positively ordered equilateral triangles $A_1A_3A_{13}$, $A_2A_1A_{21}$, and $A_3A_2A_{32}$ the coördinates of the three vertices are $A_{13}(-\omega a_1 - \omega^2 a_3)$, $A_{21}(-\omega a_2 - \omega^2 a_1)$, and $A_{32}(-\omega a_3 - \omega^2 a_2)$ where $\omega^3 = 1$. The vector $A_{32}A_1$ is the Lagrange resolvent $a_1 + \omega^2 a_2 + \omega a_3$, which we shall term u_2. Similarly, the vectors $A_{13}A_2$ and $A_{21}A_3$ are ωu_2 and $\omega^2 u_2$. Hence, we have the well-known theorem that the Lagrange resolvents $A_{32}A_1$, $A_{13}A_2$ and $A_{21}A_3$ are equal in length and intersect at angles of $2\pi/3$. These vectors meet at a point f_2 whose coördinate is

$$(2.1) \qquad f_2 = g - u_2\bar{u}_1/3\bar{u}_2. \qquad g = (a_1 + a_2 + a_3)/3$$

Similarly, if we construct on the sides of the triangle $A_1A_2A_3$ the positively-ordered equilateral triangles $A_1A_2A_{12}$, $A_2A_3A_{23}$, and $A_3A_1A_{31}$ then the vectors $A_{23}A_1$, $A_{31}A_2$ and $A_{12}A_3$ are respectively u_1, ωu_1 and $\omega^2 u_1$ where u_1 is the Lagrange resolvent $a_1 + \omega a_2 + \omega^2 a_3$. These vectors meet at the point f_1 whose coördinate is

$$(2.2) \qquad f_1 = g - u_1\bar{u}_2/3\bar{u}_1.$$

The points f_1 and f_2 are variously known as the Fermat points [7] of the triangle $A_1A_2A_3$ or as the isogenic centers.[8]

The area of the triangle $A_{13}A_{21}A_{32}$ is five-halves that of $A_1A_2A_3$ plus $3^{1/2}s^2/8$, while the area of $A_{31}A_{12}A_{23}$ is five-halves that of $A_1A_2A_3$ minus $3^{1/2}s^2/8$ where $s^2 = \overline{A_1A_2}^2 + \overline{A_2A_3}^2 + \overline{A_3A_1}^2$. Hence, the sum of the areas of the triangles $A_{13}A_{21}A_{32}$ and $A_{31}A_{12}A_{23}$ is 5 times the area of the triangle $A_1A_2A_3$. The hexagons $A_{12}A_{21}A_{32}A_{13}A_{31}A_{23}$, $A_{12}A_{31}A_{13}A_{21}A_{32}A_{23}$ and $A_{12}A_{31}A_{23}A_{32}A_{13}A_{21}$ are n-gons of type M, i. e. hexagons for which the Lagrange resolvent V_1 vanishes. Their areas are respectively $(\overline{A_1A_2}^2 + \overline{A_3A_1}^2)3^{1/2}/4$, $(\overline{A_2A_3}^2 + \overline{A_3A_1}^2)3^{1/2}/4$ and $(\overline{A_1A_2}^2 + \overline{A_2A_3}^2)3^{1/2}/4$. In terms of their five Lagrange resolvents these hexagons can be characterized respectively as $V_1 = 2V_2 + \omega V_4 = V_3 - 2\omega V_5 = 0$; $V_1 = 2V_2 + V_4 = V_3 - 2V_5 = 0$; $V_1 = 2V_2 + \omega^2 V_4 = V_3 - 2\omega^2 V_5 = 0$.

The hexagon $A_{23}A_{21}A_{31}A_{32}A_{12}A_{13}$ is worth some attention. Its area is twice that of the triangle $A_1A_2A_3$ and in terms of its resolvents we find $V_3 = V_1 - 3V_4 = 3V_2 + V_5 = 0$. To discover the geometrical significance of these conditions it is essential to express the resolvents of the hexagon in

[7] Morley, *loc. cit.*, p. 207.

[8] R. A. Johnson, *Modern Geometry*, p. 218.

terms of the resolvents of its two component triangles. Thus, if we denote by u_1 and u_2 the two Lagrange resolvents of the triangle $A_{23}A_{31}A_{12}$, and by u'_1 and u'_2 those for the triangle $A_{32}A_{13}A_{21}$ we can easily prove that *the necessary and sufficient conditions for a positively-ordered hexagon to have $V_3 = V_1 - 3V_4 = 3V_2 + V_5 = 0$ are that the centroids of its component triangles coincide, that the vector u'_1 be negatively parallel to u_1 and half its length, and that the vector u'_2 be negatively parallel to u_2 and twice its length.*

If we denote by f_1, f_2; f'_1, f'_2; F_1, F_2 the Fermat points of the triangles $A_{23}A_{31}A_{12}$, $A_{32}A_{13}A_{21}$, $A_1A_2A_3$ respectively; and by h_1, h_2; h'_1, h'_2; H_1, H_2 the Hessian points of the same triangles, then the following facts can be verified—the three triangles have the same centroid; F_1 coincides with f_2 and F_2 with f'_1; g, f_1, f'_1, h_2, h'_2 and H_1 lie on a line and so do g, f_2, f'_2, h_1, h'_1 and H_2. The distances between these points can be readily read from the relations

$$(2.3)\qquad \begin{aligned} g - f'_2 &= 2(f_2 - g); & g - f_1 &= 2(f'_1 - g) \\ g - H_1 &= 2(h_2 - g); & g - H_2 &= 2(h'_1 - g) \\ g - h_1 &= 4(g - H_2); & g - h'_2 &= 4(g - H_1). \end{aligned}$$

3. In this section, let us consider, in terms of their Lagrange resolvents, some special ordered six-points which possess features of interest. We shall first prove the theorem that

The necessary and sufficient condition for a positively-ordered hexagon to have $V_1 = V_2 = 0$ is that the sides of the triangle $x_4x_6x_2$ form positive right angles with the corresponding medians of the triangle $x_1x_3x_5$ and equal $2.3^{-1/2}$ times their length.

Since $V_1 = V_2 = 0$, we have

$$\begin{aligned} x_1 + \omega x_3 + \omega^2 x_5 &= x_4 + \omega x_6 + \omega^2 x_2 \\ x_1 + \omega^2 x_3 + \omega x_5 &= -x_4 - \omega^2 x_6 - \omega x_2 \end{aligned}$$

whence by addition,

$$3x_1 - 3g = 3^{1/2} i(x_6 - x_2). \qquad 3g = x_1 + x_3 + x_5$$

$$(3.1)\qquad \text{or}\qquad x_2 - x_6 = 3^{1/2} i(x_1 - g).$$

Similarly, we can show that

$$(3.2)\qquad x_4 - x_2 = 3^{1/2} i(x_3 - g)$$

and

$$x_6 - x_4 = 3^{1/2} i(x_5 - g)$$

which demonstrates the theorem. The conditions can easily be shown to be sufficient. Now if g' be the centroid of the triangle $x_4x_6x_2$ one can show that

$$(3.3)\qquad \begin{aligned} x_3 - x_1 &= 3^{1/2} i (x_2 - g') \\ x_5 - x_3 &= 3^{1/2} i (x_4 - g') \\ x_1 - x_5 &= 3^{1/2} i (x_6 - g') \end{aligned}$$

so that the sides of the triangle $x_1x_3x_5$ are perpendicular to the corresponding medians of the triangle $x_4x_6x_2$ and equal to $2.3^{-1/2}$ times their length. Thus, we have a mutual relationship between the two triangles.[9] In addition, since

$$\begin{vmatrix} x_4 & x_6 & x_2 \\ \bar{x}_4 & \bar{x}_6 & \bar{x}_2 \\ 1 & 1 & 1 \end{vmatrix} = 3 \begin{vmatrix} x_1 & x_3 & g \\ \bar{x}_1 & \bar{x}_3 & \bar{g} \\ 1 & 1 & 1 \end{vmatrix}$$

we see that the area of the triangle $x_4x_6x_2$ is equivalent to that of $x_1x_3x_5$. Also

$$\begin{aligned} f_2 - g &= f'_2 - g', & f_1 - g &= g' - f'_1 \\ h_1 - g &= h'_1 - g', & h_2 - g &= g' - h'_2 \end{aligned}$$

whence the vector $f_2 - g$ is equal and positively parallel to $f'_2 - g'$, but $f_1 - g$ is equal and negatively parallel to $f'_1 - g'$; etc.

From the formulae (3.1), (3.2), and (3.3), we read that *if perpendiculars, dropped from the vertices x_4, x_6, x_2 of a triangle to the corresponding sides of the triangle $x_1x_3x_5$, should meet at the centroid of $x_4x_6x_2$, then the perpendiculars from the vertices x_1, x_3, x_5 to the sides of $x_4x_6x_2$ will meet at the centroid of $x_1x_3x_5$.*

The necessary and sufficient condition for a positively-ordered six-point to have $V_4 = V_5 = 0$ is that the sides of the triangle $x_4x_6x_2$ form negative right angles with the corresponding medians of the triangle $x_1x_3x_5$ and equal $2.3^{-1/2}$ times their length. This relationship is mutual and again both triangles are equivalent in area. If in addition $V_3 = 0$, both centroids coincide, and both Fermat points f_1 and f'_1, hence the diagonals of the hexagon meet at angles of $2\pi/3$.

The necessary and sufficient condition for a positively-ordered six-point to have $V_2 = V_4 = 0$ is that the midpoint of each of its diagonals should be the midpoint of the centroids of the triangles $x_1x_3x_5$ and $x_4x_6x_2$. The two triangles have their corresponding sides equal and parallel; they are inversely equivalent in area and also perspective. The opposite sides of the hexagon are equal and negatively parallel. Also

$$\begin{aligned} f'_2 - g' &= g - f_2; & f'_1 - g' &= g - f_1 \\ h_1 - g &= g - h'_1; & h_2 - g &= g - h'_2. \end{aligned}$$

[9] If $V_1 = V_2 = V_3 = 0$, we have the special case of the above, in which the centroids g and g' coincide. See Morley, *loc. cit.*, p. 214 for details.

Hence, the join of f'_2 and f'_1 is parallel to f_1 and f_2; also the join of h'_2 and h'_1 to that of h_1 and h_2. If in addition $V_3 = 0$, we can construct the component triangles as follows—starting with the triangle $x_4x_6x_2$ with centroid g', then x_1 lies on the median from x_4 such that $\overline{x_4x_1} = 2\overline{x_4g'}$; similarly for x_3 and x_5.

The necessary and sufficient condition for a positively ordered six-point to have $V_1 = V_5 = 0$ is that each diagonal shall be parallel and equal in length to the vector joining the centroids g and g' of the two component triangles. These two triangles have their corresponding sides equal and parallel and are directly congruent. If in addition, we make $V_3 = 0$, then the two triangles will coincide throughout, and the hexagon is a doubly-counted triangle.

4. Let A_k $(k = 1, 2, \cdots, 6)$ be any positively-ordered hexagon and let us construct the following six positive hexagons for which the Lagrange resolvent V_1 vanishes, $P'_1A_2A_3A_4A_5A_6$, $P'_2A_3A_4A_5A_6A_1$, $P'_3A_4A_5A_6A_1A_2$, $P'_4A_5A_6A_1A_2A_3$, $P'_5A_6A_1A_2A_3A_4$ and $P'_6A_1A_2A_3A_4A_5$. The coordinates of the points P'_i $(i = 1, 2, \cdots, 6)$ are respectively $a_1 - V_1$, $\omega V_1 + a_2$, $a_3 - \omega^2 V_1$, $V_1 + a_4$, $a_5 - \omega V_1$ and $\omega^2 V_1 + a_6$. The equations of the six lines $A_iP'_i$ are

$$
\begin{aligned}
&\bar{V}_1 x - V_1\bar{x} + V_1\bar{a}_1 - a_1\bar{V}_1 = 0\\
&\bar{V}_1 x - \omega^2 V_1\bar{x} + \omega^2 V_1\bar{a}_2 - a_2\bar{V}_1 = 0\\
&\bar{V}_1 x - \omega V_1\bar{x} + \omega V_1\bar{a}_3 - a_3\bar{V}_1 = 0\\
&\bar{V}_1 x - V_1\bar{x} + V_1\bar{a}_4 - a_4\bar{V}_1 = 0\\
&\bar{V}_1 x - \omega^2 V_1\bar{x} + \omega^2 V_1\bar{a}_5 - a_5\bar{V}_1 = 0\\
&\bar{V}_1 x - \omega V_1\bar{x} + \omega V_1\bar{a}_6 - a_6\bar{V}_1 = 0.
\end{aligned}
$$

From the form of these equations, they represent three pairs of parallel lines, also each line makes a positively-directed angle of $2\pi/3$ with the consecutive line. The coordinates of the point of intersection of each line with the consecutive line are

$$
\begin{array}{ll}
P_1: & V_1(\bar{a}_6 - \bar{a}_1) + \bar{V}_1(a_1 - \omega^2 a_6) \div (1 - \omega^2)\bar{V}_1\\
P_2: & \omega^2 V_1(\bar{a}_1 - \bar{a}_2) + \bar{V}_1(a_2 - \omega^2 a_1) \div \quad \text{“}\\
P_3: & \omega V_1(\bar{a}_2 - \bar{a}_3) + \bar{V}_1(a_3 - \omega^2 a_2) \div \quad \text{“}\\
P_4: & V_1(\bar{a}_3 - \bar{a}_4) + \bar{V}_1(a_4 - \omega^2 a_3) \div \quad \text{“}\\
P_5: & \omega^2 V_1(\bar{a}_4 - \bar{a}_5) + \bar{V}_1(a_5 - \omega^2 a_4) \div \quad \text{“}\\
P_6: & \omega V_1(\bar{a}_5 - \bar{a}_6) + \bar{V}_1(a_6 - \omega^2 a_5) \div \quad \text{“} \quad .
\end{array}
$$

If we call the join of the lines $\overline{A_1P'_1}$ and $\overline{A_3P'_3}$ by B_5; of $\overline{A_3P'_3}$ and $\overline{A_5P'_5}$ by B_1; of $\overline{A_5P'_5}$ and $\overline{A_1P'_1}$ by B_3; of $\overline{A_6P'_6}$ and $\overline{A_2P'_2}$ by B_4; of $\overline{A_2P'_2}$ and $\overline{A_4P'_4}$ by B_6; of $\overline{A_4P'_4}$ and $\overline{A_6P'_6}$ by B_2 then one can show that $B_1B_5B_3$ and

$B_4B_2B_6$ are equal positive equilateral triangles with corresponding sides positively parallel. Now the necessary and sufficient condition that the six points of intersection of two parallel equilateral triangles—sides produced if necessary—lie on a rectangular hyperbola [10] is that the sides of the two triangles be equal. Consequently, the six points of intersection of the triangles $B_1B_5B_3$ and $B_4B_2B_6$, which are the six points P_i $(i = 1, 2, \cdots, 6)$, lie on a rectangular hyperbola. In terms of the Lagrange resolvents this six-point is characterized by $V_1 = V_3 = V_5\bar{V}_5 - 4V_3\bar{V}_3 = 0$. Hence, the sides of the triangle $P_4P_6P_2$ make positive right angles with the corresponding medians of $P_1P_3P_5$ and are equal to $2.3^{-1/2}$ times their length; also the Lagrange resolvent u_2 of the triangle $P_1P_3P_5$ is three times the length of the join of the two centroids.

Again, if A_k be any positively-ordered hexagon and we construct the six positive hexagons $P'_1A_2A_3A_4A_5A_6, \cdots$ for which the resolvent V_5 vanishes, we will obtain by a process similar to the above-mentioned one a six-point P_i for which $V_4 = V_5 = V_1\bar{V}_1 - 4V_3\bar{V}_3 = 0$. Hence, the sides of the triangle $P_4P_6P_2$ form negative right angles with the corresponding medians of $P_1P_3P_5$ and are equal to $2.3^{-1/2}$ times their length, also the Lagrange resolvent u_1 of the triangle $P_1P_3P_5$ is three times the length of the join of the two centroids. The points $B_1B_3B_5$ and $B_4B_6B_2$ are equal positive equilateral triangles with corresponding sides positively parallel and therefore the vertices of this hexagon lie on a rectangular hyperbola. Hence, associated with any positively-ordered hexagon are two circumscribed six-points, whose opposite sides are parallel and whose vertices lie on rectangular hyperbolas.

However, if we construct the six positive hexagons $P'_1A_2A_3A_4A_5A_6, \cdots$ for which the resolvent V_2 vanishes, we will obtain a six-point P_i for which $V_2 = V_4 = V_1\bar{V}_1 - 4V_3\bar{V}_3 = 0$. The points $B_1B_3B_5$ and $B_4B_6B_2$ are equal positive equilateral triangles with corresponding sides negatively parallel. Similarly, if we construct the six positive hexagons for which the resolvent V_4 vanishes, we will obtain a six-point P_i for which $V_2 = V_4 = V_5\bar{V}_5 - 4V_3\bar{V}_3 = 0$. The points $B_1B_5B_3$ and $B_4B_2B_6$ are equal positive equilateral triangles with corresponding sides negatively parallel. Hence, associated with any positively-ordered hexagon are two circumscribed six-points whose opposite sides are equal and parallel, whose diagonals pass through a point, and whose vertices lie on conics.

WESTERN RESERVE UNIVERSITY.

---

[10] J. R. Musselman, *The American Mathematical Monthly*, vol. 41 (1934), p. 634.

ON THE ABSTRACT PROPERTIES OF LINEAR DEPENDENCE.[1]

By Hassler Whitney.

1. Introduction. Let $C_1, C_2, \cdots, C_n$ be the columns of a matrix $\boldsymbol{M}$. Any subset of these columns is either linearly independent or linearly dependent; the subsets thus fall into two classes. These classes are not arbitrary; for instance, the two following theorems must hold:

(a) Any subset of an independent set is independent.

(b) If $\boldsymbol{N}_p$ and $\boldsymbol{N}_{p+1}$ are independent sets of p and $p+1$ columns respectively, then $\boldsymbol{N}_p$ together with some column of $\boldsymbol{N}_{p+1}$ forms an independent set of $p+1$ columns.

There are other theorems not deducible from these; for in § 16 we give an example of a system satisfying these two theorems but not representing any matrix. Further theorems seem, however, to be quite difficult to find. Let us call a system obeying (a) and (b) a "matroid." The present paper is devoted to a study of the elementary properties of matroids. The fundamental question of completely characterizing systems which represent matrices is left unsolved. In place of the columns of a matrix we may equally well consider points or vectors in a Euclidean space, or polynomials, etc.

This paper has a close connection with a paper by the author on linear graphs;[2] we say a subgraph of a graph is independent if it contains no circuit. Although graphs are, abstractly, a very small subclass of the class of matroids, (see the appendix), many of the simpler theorems on graphs, especially on non-separable and dual graphs, apply also to matroids. For this reason, we carry over various terms in the theory of graphs to the present theory. Remarkably enough, for matroids representing matrices, dual matroids have a simple geometrical interpretation quite different from that in the case of graphs (see § 13).

The contents of the paper are as follows: In Part I, definitions of matroids in terms of the concepts rank, independence, bases, and circuits are considered, and their equivalence shown. Some common theorems are deduced (for instance Theorem 8). Non-separable and dual matroids are studied in

[1] Presented to the American Mathematical Society, September, 1934.

[2] "Non-separable and planar graphs," *Transactions of the American Mathematical Society*, vol. 34 (1932), pp. 339-362. We refer to this paper as G.

Part II; this section might replace much of the author's paper G. The subject of Part III is the relation between matroids and matrices. In the appendix, we completely solve the problem of characterizing matrices of integers modulo 2, of interest in topology.

I. Matroids.

2. Definitions in terms of rank. Let a set M of elements $e_1, e_2, \cdots, e_n$ be given. Corresponding to each subset N of these elements let there be a number $r(N)$, the *rank* of N. If the three following postulates are satisfied, we shall call this system a *matroid.*

(R_1) *The rank of the null subset is zero.*

(R_2) *For any subset N and any element e not in N,*

$$r(N + e) = r(N) + k, \qquad (k = 0 \text{ or } 1).$$

(R_3) *For any subset N and elements e_1, e_2 not in N, if $r(N + e_1) = r(N + e_2) = r(N)$, then $r(N + e_1 + e_2) = r(N)$.*

Evidently *any subset of a matroid is a matroid.* In what follows, M is a fixed matroid. We make the following definitions:

$$\rho(N) = \text{number of elements in } N.$$

$$n(N) = \rho(N) - r(N) = \textit{nullity} \text{ of } N.$$

N is *independent,* or, the elements of N are independent, if $n(N) = 0$; otherwise, N, and its set of elements, are *dependent.*

Lemma 1. *For any N, $r(N) \geqq 0$ and $n(N) \geqq 0$. If $N \subset M$, then $r(N) \leqq r(M)$, $n(N) \leqq n(M)$.*

Lemma 2. *Any subset of an independent set is independent.*

e is *dependent on N* if $r(N + e) = r(N)$; otherwise e is *independent of N.*

A *base* is a maximal independent submatroid of M, i. e. a matroid B in M such that $n(B) = 0$, while $B \subset N$, $B \neq N$ implies $n(N) > 0$. See also Theorem 7. A *base complement* $A = M - B$ is the complement in M of a base B. A *circuit* is a minimal dependent matroid, i. e. a matroid P such that $n(P) > 0$, while $N \subset P$, $N \neq P$ implies $n(N) = 0$.[3]

Theorem 1. *N is independent if and only if it is contained in a base, or, if and only if it contains no circuit.*

[3] Compare G, Theorem 9.

THEOREM 2. *A circuit is a minimal submatroid contained in no base, i. e. containing at least one element from each base complement. A base is a maximal submatroid containing no circuit. A base complement is a minimal submatroid containing at least one element from each circuit.*

The above facts follow at once from the definitions. Note the reciprocal relationship between circuits and base complements. Note also that the definitions of independence and of being a circuit depend only on the given subset, while the property of being a base depends on the relationship of the subset to M.

3. Properties of rank. Our object here is to prove Theorem 3. The following definition will be useful:

$$\Delta(M, N) = r(M + N) - r(M). \tag{3.1}$$

LEMMA 3. $\Delta(M + e_2, e_1) \leqq \Delta(M, e_1)$.

Suppose first $r(M + e_1) = r(M) + 1$; then $r(M + e_1 + e_2) = r(M) + k$, $k = 1$ or 2. If $k = 2$, then $r(M + e_2) = r(M) + 1$, on account of (R_2), and the inequality holds; if $k = 1$, $r(M + e_2) = r(M) + l$, $l = 0$ or 1, and it holds again. If $r(M + e_2) = r(M) + 1$, the same reasoning applies. If finally $r(M + e_1) = r(M + e_2) = r(M)$, the inequality follows from (R_3).

LEMMA 4. $\Delta(M + N, e) \leqq \Delta(M, e)$.

If $N = e_1 + \cdots + e_p$, the last lemma gives

$$\Delta(M + N, e) \leqq \Delta(M + e_1 + \cdots + e_{p-1}, e) \leqq \cdots \leqq \Delta(M, e).$$

THEOREM 3. $\Delta(M + N_2, N_1) \leqq \Delta(M, N_1)$, *or,*

$$r(M + N_1 + N_2) \leqq r(M + N_1) + r(M + N_2) - r(M). \tag{3.2}$$

This is true if N_1 contains but a single element. For the general case, we apply the last lemma and induction, setting $N_1 = N' + e$:

$$\begin{aligned} \Delta(M + N_2, N_1) &= \Delta(M + N_2 + e, N') + \Delta(M + N_2, e) \\ &\leqq \Delta(M + e, N') + \Delta(M, e) = \Delta(M, N_1). \end{aligned}$$

(3.2) is evidently equivalent to:

$$r(M_1 + M_2) \leqq r(M_1) + r(M_2) - r(M_1 M_2). \tag{3.3}$$

4. Deduction of (I_1), (I_2) **from** (R_1), (R_2), (R_3). The first postulate

4

on independent sets below obviously holds if (R_1) and (R_2) hold. To prove (I_2), take N, N' as given there; then

$$r(N) = p, \qquad r(N') = p + 1.$$

We must show that for some i, $\Delta(N, e'_i) = 1$. (Then e'_i does not lie in N.) If this is not so, then on using Lemma 4 we find

$$\begin{aligned} 1 = r(N') - r(N) &\leqq \Delta(N, N') \\ &= \Delta(N, e'_1) + \Delta(N + e'_1, e'_2) + \cdots + \Delta(N + e'_1 + \cdots + e'_p, e'_{p+1}) \\ &\leqq \Delta(N, e'_1) + \Delta(N, e'_2) + \cdots + \Delta(N, e'_{p+1}) = 0, \end{aligned}$$

a contradiction.

5. Deduction of (C_1), (C_2) from (R_1), (R_2), (R_3). We shall need here a theorem showing how the nullity (or rank) of a matroid may be determined when we know what circuits it contains.

Lemma 5. *Each element of a circuit is dependent on the rest of the circuit.*

If e is an element of the circuit P, then $n(P) = 1$, $n(P - e) = 0$; hence $r(P) = \rho(P) - 1 = \rho(P - e) = r(P - e)$.

Lemma 6. *If e is dependent on P_1 but on no proper subset of P_1, then $P = P_1 + e$ is a circuit.*

As $\Delta(P_1, e) = 0$, $r(P) = r(P_1) \leqq \rho(P_1) < \rho(P)$, $n(P) > 0$, and P contains a circuit P'. If P' does not contain e, take e' in P'; then

$$\Delta(P_1 - e', e') \leqq \Delta(P' - e', e') = 0,$$

hence $r(P_1 - e') = r(P_1)$, and

$$\begin{aligned} \Delta(P_1 - e', e) &= r(P_1 - e' + e) - r(P_1 - e') \\ &\leqq r(P_1 + e) - r(P_1) = \Delta(P_1, e) = 0, \end{aligned}$$

and e is dependent on the proper subset $P_1 - e'$ of P_1, a contradiction. Therefore P' contains e. As P' is a circuit, e is dependent on the rest of P'; hence $P' = P$.

Theorem 4. *If e is not in N, there is a circuit in $N + e$ which contains e if and only if e is dependent on N.*

Suppose $P_1 + e = P$ is a circuit, $P_1 \subset N$. Then

$$\Delta(N, e) \leqq \Delta(P_1, e) = 0,$$

and e is dependent on N. Suppose, conversely, $\Delta(N, e) = 0$. Let P_1 be a smallest subset of N on which e is dependent; then by the last lemma, $P = P_1 + e$ is a circuit. (It may be that $P = e$.)

THEOREM 5. *If N is formed element by element, then $n(N)$ is just the number of times that adding an element increases the number of circuits present.*

Say $N = e_1 + \cdots + e_p$. Then if O is the null set,

$$r(N) = \Delta(O, e_1) + \Delta(e_1, e_2) + \cdots + \Delta(e_1 + \cdots + e_{p-1}, e_p).$$

Each $\Delta(e_1 + \cdots + e_{i-1}, e_i) = 0$ or 1, and $= 0$ if and only if e_i is dependent on $e_1 + \cdots + e_{i-1}$, i. e. if and only if there is a circuit in $e_1 + \cdots + e_i$ containing e_i. The number of terms is $p = \rho(N)$, and the theorem follows.

We turn now to the proof of (C_1) and (C_2). The first is obvious. To prove the second, take P_1, P_2, e_1, e_2 as given. As

$$\Delta(P_1 - e_2, e_2) = \Delta(P_2 - e_1, e_1) = 0,$$

we have

$$\Delta(P_1 + P_2 - e_2, e_2) = \Delta(P_1 + P_2 - e_1 - e_2, e_1) = 0.$$

These equations give

$$r(P_1 + P_2 - e_1 - e_2) = r(P_1 + P_2 - e_2) = r(P_1 + P_2).$$

Using (R_2) gives

$$r(P_1 + P_2 - e_1) = r(P_1 + P_2 - e_1 - e_2);$$

hence the required circuit P_3 exists, by Theorem 4.

6. Postulates for independent sets. Let M be a set of elements. Let any subset N of M be either "independent" or "dependent." Let the two following postulates be satisfied:

(I_1) *Any subset of an independent set is independent.*

(I_2) *If $N = e_1 + \cdots + e_p$ and $N' = e'_1 + \cdots + e'_{p+1}$ are independent, then for some i such that e'_i is not in N, $N + e'_i$ is independent.*

The resulting system is equivalent to a matroid, as we now show. Given any subset N of M, we let $r(N)$ be the number of elements in a largest independent subset of N. Obviously Postulates (R_1) and (R_2) are satisfied; we must prove (R_3). Say

$$r(N+e_1)=r(N+e_2)=r(N)=r.$$

Then $r(N+e_1+e_2)=r$ or $r+1$. If it equals $r+1$, there is an independent set $N'=e'_1+\cdots+e'_{r+1}$ in $N+e_1+e_2$. Let $N''=e_1''+\cdots+e_r''$ be an independent set in N. By (I_2) there is an i such that $N''+e'_i$ is an independent set of $r+1$ elements. But $N''+e'_i$ lies in $N+e_1$ or in $N+e_2$, and hence $r(N+e_1)$ or $r(N+e_2)\geqq r+1$, a contradiction. Therefore $r(N+e_1+e_2)=r$, as required.

We have shown how to deduce either set of postulates (R) or (I) from the other. Moreover the definitions of the rank and the independence or dependence of any subset of M agree under the two systems, and hence they are equivalent.

7. Postulates for bases. Let M be a set of elements, and let each subset either be or not be a "base." We assume

(B_1) *No proper subset of a base is a base.*

(B_2) *If B and B' are bases and e is an element of B, then for some element e' in B', $B-e+e'$ is a base.*

We shall prove the equivalence of this system with the preceding one. We write here $e_1e_2\cdots$ instead of $e_1+e_2+\cdots$ for short.

THEOREM 6. *All bases contain the same number of elements.*

For suppose

$$\begin{aligned} B&=e_1\cdots e_pe_{p+1}\cdots e_qe_{q+1}\cdots e_r,\\ B'&=e_1\cdots e_pe'_{p+1}\cdots e'_q \end{aligned}$$

are bases, with exactly $e_1,\cdots,e_p$ in common, and $r>q$. We might have $p=0$. $q>p$, on account of (B_1). By (B_2), we can replace e_{p+1} in B by an element e' of B', giving a base B_1. $e'=e'_{i_1}$ is one of the elements $e'_{p+1},\cdots,e'_q$, for otherwise B_1 would be a proper subset of B. Hence

$$B_1=e_1\cdots e_pe'_{i_1}e_{p+2}\cdots e_qe_{q+1}\cdots e_r.$$

If $q>p+1$, we replace e_{p+2} in B_1 by an element e'_{i_2} of B', giving a base B_2. Continuing in this manner, we obtain finally the base

$$B_{q-p} = e_1 \cdots e_p e'_{p+1} \cdots e'_q e_{q+1} \cdots e_r.$$

But this contains B' as a proper subset, contradicting (B_1).

We shall say a subset of M is independent if it is contained in a base. (I_1) obviously holds; we shall prove (I_2). Let N, N' be independent sets in the bases B, B'. Say

$$B = e_1 \cdots e_p e_{p+1} \cdots e_q e_{q+1} \cdots e_r e_{r+1} \cdots e_s,$$
$$B' = e_1 \cdots e_p e'_{p+1} \cdots e'_q e'_{q+1} \cdots e'_r e_{r+1} \cdots e_s,$$
$$N = e_1 \cdots e_p e_{p+1} \cdots e_q, \qquad N' = e_1 \cdots e_p e'_{p+1} \cdots e'_q e'_{q+1}.$$

Then N and N' have just $e_1, \cdots, e_p$ in common, and B and B' have just these elements and $e_{r+1}, \cdots, e_s$ in common. By (B_2), there is an element e'_{i_1} of B' such that

$$B_1 = B - e_{q+1} + e'_{i_1}$$

is a base. (This element cannot be any of $e_1, \cdots, e_p, e_{r+1}, \cdots, e_s$, by (B_1).) If i_1 is one of the numbers $p+1, p+2, \cdots, q+1$, then $N + e'_{i_1}$ is in a base B_1, as required. Suppose not; then there is a base

$$B_2 = B_1 - e_{q+2} + e'_{i_2},$$

with $i_2 \neq i_1$. If $p + 1 \leqq i_2 \leqq q + 1$, $N + e'_{i_2}$ is in a base B_2. If not, we find a base B_3, etc. We can drop out each of the $r - q$ elements $e_{q+1}, \cdots, e_r$ in turn; as there are only $r - q - 1$ elements e'_i with $i > q + 1$, we find at some point a base containing $e_1, \cdots, e_q, e'_j$ with $p + 1 \leqq j \leqq q + 1$. Then e'_j is in N', and $N + e'_j$ is in a base and is thus independent, as required.

The definitions of base and independent sets in the two systems (I) and (B) are easily seen to agree. Suppose (I_1) and (I_2) hold. (B_1) obviously holds; using (I_2), we prove that all bases contain the same number of elements; (B_2) now follows at once from (I_2). Hence the two systems are equivalent.

THEOREM 7. *B is a base in M if and only if*

$$r(B) = r(M), \qquad n(B) = 0.$$

Evidently B is a base under the given conditions. To prove the converse, we note first that there exists a base with $r(M)$ elements, as $r(M)$ is the maximum number of independent elements in M (see § 6). By Theorem 6, all bases have this many elements, and the equations follow.

THEOREM 8. *If B is a base and N is independent, then for some N' in B, $N + N'$ is a base.*

This follows from repeated application of Postulate (I_2) and the last theorem.

8. Postulates for circuits. Let M be a set of elements, and let each subset either be or not be a "circuit." We assume:

(C_1) *No proper subset of a circuit is a circuit.*

(C_2) *If P_1 and P_2 are circuits, e_1 is in both P_1 and P_2, and e_2 is in P_1 but not in P_2, then there is a circuit P_3 in $P_1 + P_2$ containing e_2 but not e_1.*

(C_2) may be phrased as follows: If the circuits P_1 and P_2 have the common element e, then $P_1 + P_2 - e$ is the union of a set of circuits.

We shall define the rank of any subset of M, and shall then show that the postulates for rank are satisfied. Let $e_1, \cdots, e_p$ be any ordered set of elements of M. Set $\Gamma_i = 0$ if there is a circuit in $e_1 + \cdots + e_i$ containing e_i, and set $\Gamma_i = 1$ otherwise (compare Theorem 5). Let the "rank" of $(e_1, \cdots, e_p)$ be

$$r(e_1, \cdots, e_p) = \sum_{i=1}^{p} \Gamma_i.$$

LEMMA 7. $r(e_1, \cdots, e_{q-2}, e_{q-1}, e_q) = r(e_1, \cdots, e_{q-2}, e_q, e_{q-1})$.

To prove this, let N be the ordered set $e_1, \cdots, e_{q-2}$, and set

$$r(N) = r, \qquad r(N, e_{q-1}) = r_1, \qquad r(N, e_q) = r_2,$$
$$r(N, e_{q-1}, e_q) = r_{12}, \qquad r(N, e_q, e_{q-1}) = r_{21}.$$

CASE 1. There is no circuit in $N + e_{q-1}$ containing e_{q-1}, and none in $N + e_q$ containing e_q. Then

$$r_1 = r_2 = r + 1.$$

If there is a circuit in $N + e_{q-1} + e_q$ containing e_{q-1} and e_q, then

$$r_{12} = r_1 = r_2 = r_{21};$$

otherwise,

$$r_{12} = r_1 + 1 = r_2 + 1 = r_{21}.$$

CASE 2. There is a circuit P_2 in $N + e_{q-1}$ containing e_{q-1}, and a circuit P_1 in $N + e_{q-1} + e_q$ containing e_{q-1} and e_q. Then, by (C_2), there is a circuit P_3 in $N + e_q$ containing e_q. Hence

$$r_{12} = r_1 = r = r_2 = r_{21}.$$

CASE 3. There is a circuit P_2 as above, but no circuit P_1 as above. If there is a circuit P_3 as above, the last set of equations hold. Otherwise,

$$r_{12} = r_1 + 1 = r + 1 = r_2 = r_{21}.$$

CASE 4. There is a circuit in $N + e_q$ containing e_q. This case overlaps the two preceding ones; the proof above applies here also.

LEMMA 8. *The rank of any subset N is independent of the ordering of the elements of N.*

We saw above that interchanging the last two elements of any subset does not alter the rank; hence, evidently, interchanging any two adjacent elements leaves the rank unchanged. Any ordering of M may be obtained from any other by a number of interchanges of adjacent elements; the rank remains unchanged at each step, proving the lemma.

Postulates (R_1) and (R_2) are obviously satisfied. To prove (R_3), suppose $r(N + e_1) = r(N + e_2) = r(N)$. Then there is a circuit in $N + e_1$ containing e_1 and one in $N + e_2$ containing e_2; hence $r(N + e_1 + e_2) = r(N)$.

The definitions of rank and of circuits under the two systems (R), (C) agree, and hence the systems are equivalent.

9. Fundamental sets of circuits. The circuits $P_1, \cdots, P_q$ of a matroid M form a *fundamental set of circuits* if $q = n(M)$ and the elements $e_1, \cdots, e_n$ of M can be ordered so that P_i contains e_{n-q+i} but no e_{n-q+j} $(j > i)$. The set is *strict* if P_i contains e_{n-q+i} but no e_{n-q+j} $(0 < j < i$ or $j > i)$. These sets may be called sets *with respect to* $e_{n-q+1}, \cdots, e_n$.

THEOREM 9. *If $B = e_1 + \cdots + e_{n-q}$ is a base in $M = e_1 + \cdots + e_n$, then there is a strict fundamental set of circuits with respect to $e_{n-q+1}, \cdots, e_n$; these circuits are uniquely determined.*

As $r(B) = r(M)$, $\Delta(B, e_i) = 0$ $(i = n - q + 1, \cdots, n)$. Hence, by Theorem 4, there is a circuit P_i containing e_i and elements (possibly) of B. $P_{n-q+1}, \cdots, P_n$ is the required set. Suppose, for a given i, there were also a circuit $P'_i \neq P_i$. Then Postulate (C_2) applied to P_i and P'_i would give us a circuit P in B, which is impossible.

This theorem corresponds to the theorem that if a square submatrix N of a matrix M is non-singular, then N can be turned into the unit matrix by a linear transformation on the rows of M.

THEOREM 10. *If $P_1, \cdots, P_q$ form a fundamental set of circuits with*

respect to $e_{n-q+1}, \cdots, e_n$, *then there is a unique strict set* $P'_1, \cdots, P'_q$ *with respect to* $e_{n-q+1}, \cdots, e_n$.

Set $B = M - (e_{n-q+1} + \cdots + e_n)$. The existence of $P_1, \cdots, P_q$ shows that $r(M) = r(M - e_n) = \cdots = r(B)$. Hence $\rho(B) = n - q = r(M) = r(B)$, and B is a base, by Theorem 7. Theorem 9 now applies.

Note that a matroid is not uniquely determined by a fundamental set of circuits (but see the appendix). This is shown by the following two matroids, in each of which the first two circuits form a strict fundamental set:

M, with circuits 1234, 1256, 3456;
M', with circuits 1234, 1256, 13456, 23456.

II. Separability, Dual Matroids.

10. Separable matroids. If $M = M_1 + M_2$, then $r(M) \leqq r(M_1) + r(M_2)$, on account of (3.3). If it is possible to divide the elements of M into two groups, M_1 and M_2, each containing at least one element, such that

$$r(M) = r(M_1) + r(M_2), \tag{10.1}$$

or, which is equivalent (as M_1 and M_2 have no common elements),

$$n(M) = n(M_1) + n(M_2), \tag{10.2}$$

we shall say M is *separable*; otherwise, M is *non-separable*.[4] Any single element forms a non-separable matroid. Any maximal non-separable part of M is a *component* of M.[5]

Theorem 11. *If*

$$M = M_1 + M_2, \qquad r(M) = r(M_1) + r(M_2),$$
$$M'_1 \subset M_1, \qquad M'_2 \subset M_2, \qquad M' = M'_1 + M'_2,$$

then

$$r(M') = r(M'_1) + r(M'_2).$$

Set $M_1'' = M_1 - M_1'$, $M_2'' = M_2 - M_2'$. The relations (see Theorem 3)

$$\begin{aligned} r(M) &= \Delta(M_1 + M_2', M_2'') + \Delta(M', M_1'') + r(M') \\ &\leqq \Delta(M_2', M_2'') + \Delta(M_1', M_1'') + r(M') \\ &= r(M_2) - r(M_2') + r(M_1) - r(M_1') + r(M') \end{aligned}$$

[4] Compare G, Theorem 15.
[5] See G, § 4.

together with the fact that $r(M) = r(M_1) + r(M_2)$ show that $r(M') \geqq r(M'_1) + r(M'_2)$ and hence $r(M') = r(M'_1) + r(M'_2)$.

THEOREM 12.[6] If $M = M_1 + M_2$, $r(M) = r(M_1) + r(M_2)$, *M' is non-separable, and $M' \subset M$, then either $M' \subset M_1$ or $M' \subset M_2$.*

For suppose $M' = M'_1 + M'_2$, $M'_1 \subset M_1$, $M'_2 \subset M_2$, and M'_1 and M'_2 each contain an element. By the last theorem, $r(M') = r(M'_1) + r(M'_2)$, which cannot be.

THEOREM 13. *If M_1 and M_2 are non-separable matroids with a common element e, then $M = M_1 + M_2$ is non-separable.*

For suppose $M = M'_1 + M'_2$, $r(M) = r(M'_1) + r(M'_2)$. By the last theorem, $M_1 \subset M'_1$ or $M_1 \subset M'_2$, and $M_2 \subset M'_1$ or $M_2 \subset M'_1$; this shows that either M'_1 or M'_2 is void.

THEOREM 14. *No two distinct components of M have common elements.*

This is a consequence of the last theorem. From this follows:

THEOREM 15.[7] *Any matroid may be expressed as a sum of components in a unique manner.*

THEOREM 16.[8] *A non-separable matroid M of nullity 1 is a circuit, and conversely.*

If M_1 is a proper non-null subset of the non-separable matroid M, and $M_2 = M - M_1$, then $r(M) < r(M_1) + r(M_2)$. Hence

$$1 = n(M) > n(M_1) + n(M_2),$$

and $n(M_1) = 0$, proving that M is a circuit.

Conversely, if $M = M_1 + M_2$ is a circuit, and M_1 and M_2 each contain elements, then

$$\begin{aligned} r(M_1) + r(M_2) &= \rho(M_1) + \rho(M_2) - n(M_1) - n(M_2) \\ &= \rho(M) > r(M), \end{aligned}$$

showing that M is non-separable.

[6] Compare G, Lemma, p. 344.
[7] Compare G, Theorem 12.
[8] Compare G, Theorem 10.

LEMMA 9. *Let $M = M_1 + M_2$ be non-separable, and let M_1 and M_2 each contain elements but have no common elements. Then there is a circuit P in M containing elements of both M_1 and M_2.*

Suppose there were no such circuit. Say $M_2 = e_1 + \cdots + e_s$. Using Theorem 4, we see that

$$\Delta(M_1 + e_1 + \cdots + e_{i-1}, e_i) = \Delta(e_1 + \cdots + e_{i-1}, e_i) \quad (i = 1, \cdots, s),$$

and hence $r(M) = r(M_1) + r(M_2)$, a contradiction.

THEOREM 17.[9] *Any non-separable matroid M of nullity $n > 0$ can be built up in the following manner: Take a circuit M_1; add a set of elements which forms a circuit with one or more elements of M_1, forming a non-separable matroid M_2 of nullity 2 (if $n(M) > 1$); repeat this process till we have $M_n = M$.*

As $n > 0$, M contains a circuit M_1. If $n > 1$, we use the preceding lemma $n - 1$ times. The matroid at each step is non-separable, by Theorems 16 and 13.

THEOREM 18.[10] *Let $M = M_1 + \cdots + M_p$, and let $M_1, \cdots, M_p$ be non-separable. Then the following statements are equivalent:*

(1) *$M_1, \cdots, M_p$ are the components of M.*

(2) *No two of the matroids $M_1, \cdots, M_p$ have common elements, and there is no circuit in M containing elements of more than one of them.*

(3) *$r(M) = r(M_1) + \cdots + r(M_p)$.*

We cannot replace rank by nullity in (3); see G, p. 347.

(2) follows from (1) on application of Theorems 13 and 16.

To prove (1) from (2), take any M_i. If it is not a component of M, there is a larger non-separable submatroid M'_i of M containing it. By Lemma 9, there is a circuit P in M'_i containing elements of M_i and elements not in M_i; P must contain elements of some other M_j, a contradiction.

Next we prove (3) from (1). If $p > 1$, M is separable; say $M = M'_1 + M'_2$, $r(M) = r(M'_1) + r(M'_2)$. By Theorem 12, each M_i is in either M'_1 or M'_2; hence M'_1 and M'_2 are each a sum of components of M. If one of these

---

[9] See G, Theorem 19; also Whitney, "2-isomorphic graphs," *American Journal of Mathematics*, vol. 55 (1933), p. 247, footnote.

[10] Compare G, Theorem 17.

contains more than one component, we separate it similarly, etc. (3) now follows easily.

Finally we prove (1) from (3). Let M' be a component of M, and suppose it has an element in M_i. As

$$r(M) = r(M_i) + \sum_{j \neq i} r(M_j),$$

M' is contained in M_i, by Theorem 12; as M_i is non-separable, $M' = M_i$.

THEOREM 19.[11] *The elements e_1 and e_2 are in the same component of M if and only if they are contained in a circuit P.*

If e_1 and e_2 are both in P, they are part of a non-separable matroid, which lies in a single component of M. Suppose now e_1 and e_2 are in the same component M_0 of M, and suppose there is no circuit containing them both. Let M_1 be e_1 plus all elements which are contained in a circuit containing e_1. By Lemma 9, there is a subset M^* of $M_0 - M_1$ which forms with part of M_1 a circuit P_3. P_3 does not contain e_1. If e'_4 is an element of P_3 in M_1, there is a circuit P_1 in M_1 containing e_1 and e'_4. Let e_3 be an element of M^*. Then in $M_1 + M^*$ there are circuits P_1 and P_3 which contain e_1 and e_3 respectively, and have a common element.

Let M' be a smallest subset of M_0 which contains circuits P'_1 and P'_3 such that one contains e_1, the other contains e_3, and they have common elements. Then P'_1 and P'_3 are distinct, and $M' = P'_1 + P'_3$. Let e_4 be a common element. By Postulate (C_2), there is a circuit P_1 in $M' - e_4$ containing e_1, and a circuit P_3 in $M' - e_4$ containing e_3. By the definition of M', P_1 and P_3 have no common elements. By Postulate (C_1), P_1 is not contained in P'_1; hence it contains an element e_5 of $M' - P'_1$. P_3 does not contain e_5. As P_3 is not contained in P'_3, it contains an element e_6 of P'_1. But now P'_1 contains e_1, P_3 contains e_3, $P'_1 + P_3$ have a common element e_6, and $P'_1 + P_3$ does not contain e_5 and is thus a proper subset of M', a contradiction. This proves the theorem.

11. Dual matroids. Suppose there is a $1-1$ correspondence between the elements of the matroids M and M', such that if N is any submatroid of M and N' is the complement of the corresponding matroid of M', then

$$r(N') = r(M') - n(N). \tag{11.1}$$

[11] Compare D. König, *Acta Litterarum ac Scientiarum Szeged*, vol. 6, pp. 155-179, 4. (p. 159). The present theorem shows that a "glied" is the same as a component.

n(M)

We say then that M' is a *dual* of M.[12]

THEOREM 20. *If M' is a dual of M, then*

$$r(M') = n(M), \qquad n(M') = r(M).$$

Set $N = M$; then $n(N) = n(M)$. In this case N' is the null matroid, and $r(N') = 0$. (11.1) now gives $r(M') = n(M)$. Also

$$n(M') = \rho(M') - r(M') = \rho(M) - n(M) = r(M).$$

THEOREM 21. *If M' is a dual of M, then M is a dual of M'.*

Take any N and corresponding N' as before. The equations

$$r(N') = r(M') - n(N), \qquad r(M') = n(M),$$
$$\rho(N) + \rho(N') = \rho(M)$$

give

$$\begin{aligned} r(N) &= \rho(N) - n(N) = \rho(N) - [r(M') - r(N')] \\ &= \rho(N) - n(M) + [\rho(N') - n(N')] \\ &= \rho(M) - n(M) - n(N') = r(M) - n(N'), \end{aligned}$$

as required.

THEOREM 22. *Every matroid has a dual.*

This is in marked contrast to the case of graphs, for only a planar graph has a dual graph (see G, Theorem 29).

Let M' be a set of elements in 1 — correspondence with elements of M. If N' is any subset of M', let N be the complement of the corresponding subset of M, and set $r(N') = n(M) - n(N)$. (R_1), (R_2), (R_3) are easily seen to hold in M', as they hold in M; hence M' is a matroid. Obviously $r(M') = n(M)$, and M' is a dual of M.

THEOREM 23. *M and M' are duals if and only if there is a $1-1$ correspondence between their elements such that bases in one correspond to base complements in the other.*

Suppose first M and M' are duals. Let B be a base in either matroid, say in M, and let B' be the complement of the corresponding submatroid of the other matroid, M'. Then

---

[12] Compare G, § 8. Theorems 20, 21, 24, 25 correspond to Theorems 20, 21, 23, 25 in G. Note that *two duals of the same matroid are isomorphic*, that is, there is a $1-1$ correspondence between their elements such that corresponding subsets have the same rank. Such a statement cannot be made about graphs. Compare H. Whitney, "2-isomorphic graphs," *American Journal of Mathematics*, vol. 55 (1933), pp. 245-254.

$$r(B') = r(M') - n(B) = r(M'),$$
$$n(B') = r(M) - r(B) = 0,$$

and B' is a base in M', by Theorem 7.

Suppose, conversely, that bases in one correspond to base complements in the other. Let N be a submatroid of M and let N' be the complement of the corresponding submatroid of M'. There is a base B' in M' with $r(N')$ elements in N', by Theorem 8. The complement in M of the submatroid corresponding to B' in M' is a base B in M with $\rho(N') - r(N') = n(N')$ elements in $M - N$, and hence with $r(M) - n(N')$ elements in N. This shows that

$$r(N) = r(M) - n(N') + k, \qquad k \geqq 0.$$

In a similar fashion we see that

$$r(N') = r(M') - n(N) + k', \qquad k' \geqq 0.$$

As B contains $r(M)$ elements and B' contains $r(M')$ elements, $r(M) + r(M') = \rho(M)$. Hence, adding the above equations,

$$k + k' = r(N) + r(N') + n(N) + n(N') - r(M) - r(M')$$
$$= \rho(N) + \rho(N') - \rho(M) = 0.$$

Hence $k = 0$, and the first equation above shows that M and M' are duals.

There are various other ways of stating conditions on certain submatroids of M and M' which will ensure these matroids being duals.[13]

THEOREM 24. *Let $M_1, \cdots, M_p$ and $M'_1, \cdots, M'_p$ be the components of M and M' respectively, and let M'_i be a dual of M_i $(i = 1, \cdots, p)$. Then M' is a dual of M.*

Let N be any submatroid of M, and let the parts of N in $M_1, \cdots, M_p$ be $N_1, \cdots, N_p$. Let N'_i be the complement in M'_i of the submatroid corresponding to N_i; then $N' = N'_i + \cdots + N'_p$ is the complement in M' of the submatroid corresponding to N. By Theorems 18 and 11 we have

$$r(N') = r(N'_1) + \cdots + r(N'_p), \qquad n(N) = n(N_1) + \cdots + n(N_p).$$

Also

$$r(M') = r(M'_1) + \cdots + r(M'_p), \qquad r(N'_i) = r(M'_i) - n(N_i);$$

adding the last set of equations gives $r(N') = r(M') - n(N)$, as required.

---

[13] See for instance a paper by the author "Planar graphs," *Fundamenta Mathematicae*, vol. 21 (1933), pp. 73-84, Theorem 2. Cut sets may of course be defined in terms of rank.

THEOREM 25. *Let M and M' be duals, and let $M_1, \cdots, M_p$ be the components of M. Let $M'_1, \cdots, M'_p$ be the corresponding submatroids of M'. Then $M'_1, \cdots, M'_p$ are the components of M', and M'_i is a dual of M_i $(i = 1, \cdots, p)$.*

The complement in M of the submatroid corresponding to M'_i in M' is $\sum_{j \neq i} M_j$. Hence, as M and M' are duals and the M_j $(j \neq i)$ are the components of $\sum_{j \neq i} M_j$ (see Theorem 18),

$$r(M'_i) = r(M') - n(\sum_{j \neq i} M_j) = r(M') - \sum_{j \neq i} n(M_j).$$

Adding gives

$$\begin{aligned}\sum_i r(M'_i) &= pr(M') - (p-1) \sum_j n(M_j) = pr(M') - (p-1) n(M) \\ &= pr(M') - (p-1) r(M') = r(M').\end{aligned}$$

Therefore, by Theorem 12, each component of M' is contained in some M'_i. In the same way we see that each component of M is contained in a matroid corresponding to a component of M'; hence the components of one matroid correspond exactly to the components of the other.

Let N_i be any submatroid of M_i, and let N' and N'_i be the complements in M' and M'_i of the submatroid corresponding to N_i. The equations

$$r(M') = \sum_j r(M'_j), \qquad r(N') = r(N'_i) + \sum_{j \neq i} r(M'_j),$$
$$r(N') = r(M') - n(N_i),$$

give

$$r(N'_i) = r(M'_i) - n(N_i),$$

which shows that M'_i is a dual of M_i.

THEOREM 26. *A dual of a non-separable matroid is non-separable.*

This is a consequence of the last theorem.

III. MATRICES AND MATROIDS.

12. Matrices, matroids, and hyperplanes. Consider the matrix

$$M = \begin{Vmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} \end{Vmatrix};$$

let its columns be $C_1, \cdots, C_n$. Any subset $\boldsymbol{N}$ of these columns forms a matrix, and this matrix has a rank, $r(\boldsymbol{N})$. If we consider the columns as abstract elements, we have a matroid M. The proof of this is simple if we consider the rank of a matrix as the number of linearly independent columns in it. (R_1) and (R_2) are then obvious. To prove (R_3), suppose $r(\boldsymbol{N}+C_1) = r(\boldsymbol{N}+C_2) = r(\boldsymbol{N})$; then C_1 and C_2 can each be expressed as a linear combination of the other columns of $\boldsymbol{N}$, and hence $r(\boldsymbol{N}+C_1+C_2) = r(\boldsymbol{N})$. The terms independent and base carry over to matrices and agree with the ordinary definitions; a base in $\boldsymbol{M}$ is a minimal set of columns in terms of which all remaining columns of $\boldsymbol{M}$ may be expressed.

We may interpret $\boldsymbol{M}$ geometrically in two different ways; the second is the more interesting for our purposes:

(a) Let E_m be Euclidean space of m dimensions. Corresponding to each column C_i of $\boldsymbol{M}$ there is a point X_i in E_m with coördinates $a_{1i}, \cdots, a_{mi}$. The subset $C_{i_1}, \cdots, C_{i_p}$ of $\boldsymbol{M}$ is linearly independent if and only if the points $O = (0, \cdots, 0)$, $X_{i_1}, \cdots, X_{i_p}$ are linearly independent in E_m, i. e. if and only if these $p+1$ points determine a hyperplane in E_m of dimension p. A base in $\boldsymbol{M}$ corresponds to a minimal set of points $X_{i_1}, \cdots, X_{i_p}$ in E_m such that each X_j of $\boldsymbol{M}$ lies in the hyperplane determined by $O, X_{i_1}, \cdots, X_{i_p}$. Then p is the rank of $\boldsymbol{M}$.

(b) Let E_n be Euclidean space of n dimensions. Let $R_1, \cdots, R_m$ be the rows of $\boldsymbol{M}$. If $Y_1, \cdots, Y_m$ are the corresponding points of E_n: $Y_i = (a_{i1}, \cdots, a_{in})$, then the points $O, Y_1, \cdots, Y_m$ determine a hyperplane $H = H(\boldsymbol{M})$, which we shall call the *hyperplane associated with* $\boldsymbol{M}$. The dimension $d(H)$ of H is $r(\boldsymbol{M})$. Let $\boldsymbol{N} = C_{i_1} + \cdots + C_{i_p}$ be a subset of $\boldsymbol{M}$, and let E' be the p-dimensional coördinate subspace of E_n containing the x_{i_1} and ... and the x_{i_p} axes. The j-th row of $\boldsymbol{N}$ corresponds to the point Y'_j in E' with coördinates $(a_{ji_1}, \cdots, a_{ji_p})$; this is just the projection of Y_j onto E'. If H' is the hyperplane in E' determined by the points $O, Y'_1, \cdots, Y'_m$, then H' is exactly the projection of H onto E', and

$$d(H') = r(\boldsymbol{N}). \tag{12.1}$$

Let $\boldsymbol{N} = (C_{i_1}, \cdots, C_{i_p})$ be any subset of $\boldsymbol{M}$, and let E', H' correspond to $\boldsymbol{N}$. Then $\boldsymbol{N}$ *is independent if and only if*

$$d(H') = p,$$

and is a base if and only if

$$d(H') = d(H) = p.$$

THEOREM 27. *There is a unique matroid M associated with any hyperplane H through the origin in E_n.*

Let M contain the elements $e_1, \cdots, e_n$, one corresponding to each coördinate of E_n. Given any subset $e_{i_1}, \cdots, e_{i_p}$, we let its rank be the dimension of the projection of H onto the corresponding coördinate hyperplane E' of E_n. It was seen above that if $\boldsymbol{M}$ is any matrix determining H, then M is the matroid associated with $\boldsymbol{M}$.

13. Orthogonal hyperplanes and dual matroids. We prove the following theorem:

THEOREM 28. *Let H be a hyperplane through the origin in E_n, of dimension r, and let H' be the orthogonal hyperplane through the origin, of dimension $n - r$. Let M and M' be the associated matroids. Then M and M' are duals.*

We shall show that bases in one matroid correspond to base complements in the other; Theorem 23 then applies. Let

$$\boldsymbol{M} = \begin{Vmatrix} a_{11} \cdots a_{1n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{r1} \cdots a_{rn} \end{Vmatrix}, \qquad \boldsymbol{M}' = \begin{Vmatrix} b_{11} \quad \cdots b_{1n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ b_{n-r,1} \cdots b_{n-r,n} \end{Vmatrix}$$

be matrices determining H and H' respectively. Say the first r columns of $\boldsymbol{M}$ form a base in $\boldsymbol{M}$, i. e. the corresponding determinant A is $\neq 0$. As H and H' are orthogonal, we have for each i and j

$$a_{i1}b_{j1} + a_{i2}b_{j2} + \cdots + a_{in}b_{jn} = 0.$$

Keeping j fixed, we have a set of r linear equations in the b_{jk}. Transpose the last $n - r$ terms in each equation to the other side, and solve for b_{jk}. We find

$$b_{jk} = \frac{-1}{A} \sum_{l=r+1}^{n} b_{jl} \begin{vmatrix} a_{11} \cdots a_{1l} \cdots a_{1r} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{r1} \cdots a_{rl} \cdots a_{rr} \end{vmatrix} = \sum_{l=r+1}^{n} c_{kl} b_{jl} \qquad (k = 1, \cdots, r).$$

This is true for each $j = 1, \cdots, n - r$, and the c_{kl} are independent of j. Thus the k-th column of $\boldsymbol{M}'$ is expressed in terms of the last $n - r$ columns. As this is true for $k = 1, \cdots, r$, the last $n - r$ columns form a base in $\boldsymbol{M}'$, as required.

14. The circuit matrix of a given matrix. Consider the matrix $\boldsymbol{M}$ of § 12. Suppose the columns $C_{i_1}, \cdots, C_{i_p}$ form a circuit, i. e. the corresponding

elements of the corresponding matroid form a circuit. Then these columns are linearly dependent, and there are numbers $b_1, \cdots, b_n$ such that

$$\text{(14.1)} \qquad \begin{array}{ll} a_{i1}b_1 + \cdots + a_{in}b_n = 0 & (i = 1, \cdots, m), \\ b_j = 0 \quad (j \neq i_1, \cdots, i_p), & b_j \neq 0 \quad (j = i_1, \cdots, i_p). \end{array}$$

The b_j are all $\neq 0$ $(j = i_1, \cdots, i_p)$, for otherwise a proper subset of the columns would be dependent, contrary to the definition of a circuit. (They are uniquely determined except for a constant factor; see Lemma 11.) Suppose the circuits of $\boldsymbol{M}$ are $P_1, \cdots, P_s$. Then there are corresponding sets of numbers $b_{i1}, \cdots, b_{in}$ $(i = 1, \cdots, s)$, forming a matrix

$$\boldsymbol{M}' = \left\| \begin{array}{ccc} b_{11} & \cdots & b_{1n} \\ \cdot & \cdot & \cdot \\ b_{s1} & \cdots & b_{sn} \end{array} \right\|,$$

the *circuit matrix* of the matrix $\boldsymbol{M}$.

THEOREM 29. *Let $P_1, \cdots, P_q$ be a fundamental set of circuits in $\boldsymbol{M}$* (see § 9). *Then the corresponding rows of the circuit matrix $\boldsymbol{M}'$ form a base for the rows of $\boldsymbol{M}'$. Hence $r(\boldsymbol{M}') = q = n(\boldsymbol{M})$.*

Suppose the columns of $\boldsymbol{M}$ are ordered so that P_i contains C_{n-q+i} but no column C_{n-q+j} $(j > i)$. Then if the corresponding row of $\boldsymbol{M}'$ is $R'_i = (b_{i1}, \cdots, b_{in})$, we have $b_{i,n-q+i} \neq 0$ and $b_{i,n-q+j} = 0$ $(j > i)$. Hence the rows $R'_1, \cdots, R'_q$ of $\boldsymbol{M}'$ are linearly independent, and $r(\boldsymbol{M}') \geqq q$. Hence $r(\boldsymbol{M}') = n(\boldsymbol{M}) = q$, and each row of $\boldsymbol{M}'$ may be expressed in terms of $R'_1, \cdots, R'_q$.

THEOREM 30. *If $\boldsymbol{M}'$ is the circuit matrix of $\boldsymbol{M}$ and H', H are the corresponding hyperplanes, then H' is the hyperplane of maximum dimension orthogonal to H.*

This is a consequence of (14.1) and the last theorem.

THEOREM 31. *The matroids corresponding to a matrix and its circuit matrix are duals.*

This follows from the last theorem and Theorem 28.

15. On the structure of a circuit matrix. Let M be any matroid, and M', its dual. If there exists a matrix $\boldsymbol{M}$ corresponding to M, it is perhaps most easily constructed by considering it as the circuit matrix of a matrix $\boldsymbol{M}'$

corresponding to M'. Let H and H' be the hyperplanes corresponding to $\mathbf{M}$ and $\mathbf{M}'$. We shall say the set of numbers $(a_1, \cdots, a_n)$ is in $Z_{i_1 \ldots i_p}$ if

$$a_j \neq 0 \quad (j = i_1, \cdots, i_p), \qquad a_j = 0 \quad (j \neq i_1, \cdots, i_p).$$

If $(a_1, \cdots, a_n)$ is in H and in $Z_{i_1 \ldots i_p}$, then the columns $C_{i_1}, \cdots, C_{i_p}$ of $\mathbf{M}'$ are dependent, evidently.

LEMMA 10. *Let $(b_1, \cdots, b_n)$ be a point of H. If it is in $Z_{i_1 \ldots i_p}$, then the matroid $N' = e_{i_1} + \cdots + e_{i_p}$ is the union of a set of circuits in M'.*

Here e_i in M' corresponds to C_i in $\mathbf{M}$. We need merely show that for each i_s there is a circuit P in N' containing e_{i_s}. Let $k_1 = i_s, k_2, \cdots, k_q$ be a minimal set of numbers from $(i_1, \cdots, i_p)$ containing i_s such that there is a point $(c_1, \cdots, c_n)$ of H in $Z_{k_1 \ldots k_q}$; then $e_{k_1} + \cdots + e_{k_q}$ is the required circuit. For if it were not a circuit, there would be a proper subset $(l_1, \cdots, l_r)$ of $(k_1, \cdots, k_q)$ and a point $(d_1, \cdots, d_n)$ of H in $Z_{l_1 \ldots l_r}$. No $l_i = k_1$, on account of the minimal property of $(k_1, \cdots, k_q)$. Say $l_1 = k_t$, and set

$$a_i = d_{k_t} c_i - c_{k_t} d_i \qquad (i = 1, \cdots, n).$$

Then $(a_1, \cdots, a_n)$ is in H and in $Z_{m_1 \ldots m_u}$ with $(m_1, \cdots, m_u)$ a proper subset of $(k_1, \cdots, k_q)$ containing k_1, again a contradiction.

LEMMA 11. *If $P = e_{i_1} + \cdots + e_{i_p}$ is a circuit of M' and $(b_1, \cdots, b_n)$ and $(b'_1, \cdots, b'_n)$ are in H and in $Z_{i_1 \ldots i_p}$, then these two sets are proportional.*

For otherwise, $(c_1, \cdots, c_n)$ with $c_i = b'_{i_1} b_i - b_{i_1} b'_i$ would be a point of H in some $Z_{k_1 \ldots k_q}$ with $(k_1, \cdots, k_q)$ a proper subset of $(i_1, \cdots, i_p)$, and P would not be a circuit.

It is instructive to show directly that Postulate (C_2) holds for matrices: P_1 and P_2 are represented by rows $(b_1, \cdots, b_n)$ and $(b'_1, \cdots, b'_n)$ of $\mathbf{M}$, lying in $Z_{12 i_1 \ldots i_p}$ and $Z_{1 k_1 \ldots k_q}$ respectively, where $k_1, \cdots, k_q \neq 2$. Set $c_i = b'_1 b_i - b_1 b'_i$; then $(c_1, \cdots, c_n)$ is in H and in $Z_{2 l_1 \ldots l_r}$, with $(l_1, \cdots, l_r)$ a subset of $(i_1, \cdots, i_p, k_1, \cdots, k_q)$; the existence of P_3 now follows from Lemma 10.

THEOREM 32. *Let $\mathbf{M}$ be the circuit matrix of $\mathbf{M}'$. Let $P_1, \cdots, P_q$ form a strict fundamental set of circuits in M' with respect to $e_{n-q+1}, \cdots, e_n$, and let the first q rows in $\mathbf{M}$ correspond to $P_1, \cdots, P_q$. Let $(i_1, \cdots, i_s)$ be any set of numbers from $(1, \cdots, q)$, let $(j_1, \cdots, j_s)$ be any set from $(1, \cdots, n-q)$, and let $(i'_1, \cdots, i'_{q-s})$ be the set complementary to $(i_1, \cdots, i_s)$ in $(1, \cdots, q)$.*

Then the determinant D in $\boldsymbol{M}$ with rows $i_1, \cdots, i_s$ and columns $j_1, \cdots, j_s$ equals zero if and only if the determinant D' with rows $1, \cdots, q$ and columns $j_1, \cdots, j_s$, $n-q+i'_1, \cdots, n-q+i'_{q-s}$ equals zero, or, if and only if there exists a circuit P in $\boldsymbol{M}'$ containing none of the columns $e_{j_1}, \cdots, e_{j_s}$, $e_{n-q+i'_1}, \cdots, e_{n-q+i'_{q-s}}$.

In the matrix of the last $q = r(\boldsymbol{M})$ columns of $\boldsymbol{M}$, the terms along the main diagonal and only those are $\neq 0$. If we expand D' by Laplace's expansion in terms of the columns $n-q+i'_1, \cdots, n-q+i'_{q-s}$, we see at once that $D' = 0$ if and only if $D = 0$.

Suppose $D = 0$. Then there is a set of numbers $(\alpha_1, \cdots, \alpha_q)$, not all zero, with $\alpha_i = 0$ $(i \neq i_1, \cdots, i_s)$, such that

$$b_k = \alpha_1 b_{1k} + \cdots + \alpha_q b_{qk} = 0 \qquad (k = j_1, \cdots, j_s),$$

$(b_{i1}, \cdots, b_{in})$ being the i-th row of $\boldsymbol{M}$., $b_k = 0$ also for $k = n - q + i'_1, \cdots$, $n - q + i'_{q-s}$, as each term is zero for such k. The point $(b_1, \cdots, b_n)$ is in H. Any circuit given by Lemma 10 is the required circuit P.

Suppose the circuit P exists. Then it is represented by a row $(b_1, \cdots, b_n)$ in $\boldsymbol{M}$. As the first q rows of $\boldsymbol{M}$ are of rank $q = r(\boldsymbol{M})$, $(b_1, \cdots, b_n)$ can be expressed in terms of them; say $b_k = \Sigma \alpha_i b_{ik}$. As $b_k = 0$ $(k = n - q + i'_1, \cdots$, $n - q + i'_{q-s})$, certainly $\alpha_k = 0$ $(k = i'_1, \cdots, i'_{q-s})$. $D = 0$ now follows from the fact that $b_k = 0$ $(k = j_1, \cdots, j_s)$.

16. A matroid with no corresponding matrix.[14] The matroid M' has seven elements, which we name $1, \cdots, 7$. The bases consist of all sets of three elements except

(16.1) 124, 135, 167, 236, 257, 347, 456.

Defining rank in terms of bases, we have: Each set of k elements is of rank k if $k \leqq 2$ and of rank 3 if $k \geqq 4$; a set of three elements is of rank 2 if the set is in (16.1) and is of rank 3 otherwise. It is easy to see that the postulates for rank are satisfied. (R_3) in the case that N contains two elements is satisfied vacuously. For suppose $r(N + e_1) = r(N + e_2) = r(N) = 2$. Then $N + e_1$ and $N + e_2$ are both in (16.1); but any two of these sets have but a single element in common.

---

[14] After the author had noted that M' satisfies (C^*) and corresponds to no linear graph, and had discovered a matroid with nine elements corresponding to no matrix, Saunders MacLane found that M' corresponds to no matrix, and is a well known example of a finite projective geometry (see O. Veblen and J. W. Young, *Projective Geometry*, pp. 3-5).

If there exists a matrix $\boldsymbol{M}'$, corresponding to M', then let $\boldsymbol{M}$ be its circuit matrix. 123 is a base in M', and hence

$$(16.2) \qquad 124, \quad 135, \quad 236, \quad 1237$$

form a fundamental set of circuits in M'. Let R_1, R_2, R_3, R_4 be the corresponding rows of $\boldsymbol{M}$. By multiplying in succession row 1, column 2, rows 2, 3, 4, and columns 4, 5, 6, 7 by suitable constants $\neq 0$, we bring $\boldsymbol{M}$ into the following form:

$$(16.3) \qquad \boldsymbol{M} = \left\| \begin{array}{ccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & a & 0 & 1 & 0 & 0 \\ 0 & 1 & b & 0 & 0 & 1 & 0 \\ 1 & c & d & 0 & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\| ;$$

a, b, c and d are $\neq 0$. We now apply Theorem 32 with

$$(i_1, \cdots, i_s; j_1, \cdots, j_s) = (1,4;\ 1,2), \quad (2,4;\ 1,3), \quad (3,4;\ 2,3),$$

i. e. using the circuits 347, 257, 167. This gives

$$\begin{vmatrix} 1 & 1 \\ 1 & c \end{vmatrix} = \begin{vmatrix} 1 & a \\ 1 & d \end{vmatrix} = \begin{vmatrix} 1 & b \\ c & d \end{vmatrix} = 0,$$

and hence $c = 1$, $a = d = b$. Using the circuit 456, with sets $(1,2,3;\ 1,2,3)$ gives $2a = 0$, $a = 0$, a contradiction.

In regard to this example, see the end of the paper.

APPENDIX.

Matrices of Integers mod 2.

We wish to characterize those matroids M corresponding to matrices $\boldsymbol{M}$ of integers mod 2,[15] i. e. matrices whose elements are all 0 or 1, where rank etc. is defined mod 2. We shall consider linear combinations, *chains:*

$$(A.1) \qquad \alpha_1 e_1 + \cdots + \alpha_n e_n \qquad (\alpha\text{'s integers mod } 2)$$

in the elements of M. The α's may be taken as 0 or 1; (A.1) may then be interpreted as the submatroid N whose elements have the coefficient 1. Conversely, any $N \subset M$ may be written as a chain. Submatroids are added

[15] See O. Veblen, "Analysis situs," 2nd ed., *American Mathematical Society Colloquium Publications*, Ch. I and Appendix 2.

(mod 2) by adding the corresponding chains (mod 2). For instance, $(e_1 + e_2) + (e_2 + e_3) \equiv e_1 + e_3 \pmod 2$.

Any sum (mod 2) of circuits in M we shall call a *cycle* in M. N is the *true sum* of $N_1, \cdots, N_s$ if these latter have no common elements and $N = N_1 + \cdots + N_s$. We consider matroids which satisfy the following postulate:

(C^*) *Each cycle is a true sum of circuits.*

Postulate (C_2) is a consequence of (C^*). For the cycle $P_1 + P_2$ is a submatroid containing e_2 but not e_1; The existence of P_3 now follows from (C^*).

A simple example of a matroid not satisfying (C^*) is given by the matroid M' at the end of § 9.

THEOREM 33. *A circuit is a minimal non-null cycle, and conversely.*

This is proved with the aid of Postulates (C_1) and (C^*).

THEOREM 34. *Let $P_1, \cdots, P_q$ be a strict fundamental set of circuits in M with respect to $e_{n-q+1}, \cdots, e_n$. Then there are exactly 2^q cycles in M, formed by taking all sums* (mod 2) *of $P_1, \cdots, P_q$.*

First, each sum $P_{i_1} + \cdots + P_{i_s}$ (mod 2) is a cycle, containing $e_{n-q+i_1}, \cdots, e_{n-q+i_s}$ and elements (perhaps) from $B = e_1, \cdots, e_{n-q}$; obviously distinct sums give distinct cycles. Now let Q be any cycle in M; say Q contains $e_{n-q+k_1}, \cdots, e_{n-q+k_r}$ and elements (perhaps) from B. Set $Q' = P_{k_1} + \cdots + P_{k_r}$; then $Q + Q'$ is a cycle containing elements from B alone. But B is a base (see the proof of Theorem 10), and hence contains no circuits. Consequently $Q + Q'$ is the null cycle, and $Q = Q'$.

THEOREM 35. *As soon as the circuits of a strict fundamental set are known, all the circuits may be determined.*

This is a consequence of the last two theorems. It is to be contrasted with the final remark of § 9.

Remark. The word "strict" may be omitted in the last two theorems.

THEOREM 36. *Let $e_1, \cdots, e_n$ be a set of elements, and let $P_1, \cdots, P_q$ be any subsets such that P_i contains e_{n-q+i} and possibly elements from $e_1, \cdots, e_{n-q}$ alone. Then there is a unique matroid M satisfying* (C^*), *with $P_1, \cdots, P_q$ as a strict fundamental set of circuits.*

We form the 2^q cycles of Theorem 34. Those cycles which contain no other non-null cycle as a proper subset we call circuits; in particular, $P_1, \cdots, P_q$ are circuits. To prove (C*), let Q be a non-null cycle. If it is not a circuit, it contains a circuit P as a proper subset. Q and P are sums (mod 2) from $P_1, \cdots, P_q$, hence the same is true of $Q - P$, and $Q - P$ is one of the 2^q cycles. If it is not a circuit, we again extract a circuit, etc.

This theorem furnishes a simple method of constructing all matroids satisfying (C*).

We turn now to the study of matrices of integers (mod 2)

$$\boldsymbol{M} = \begin{Vmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} \end{Vmatrix} \qquad \text{(each } a_{ij} = 0 \text{ or } 1\text{)}.$$

Any linear combination (mod 2) of the columns

$$\text{(A. 2)} \qquad \alpha_1 C_1 + \cdots + \alpha_n C_n \qquad (\alpha\text{'s integers mod } 2)$$

is a set of numbers $(\Sigma\alpha_i a_{1i}, \cdots, \Sigma\alpha_i a_{mi})$, which we call a *chain* (mod 2) in $\boldsymbol{M}$. As before, we may take each coefficient as 0 or 1, and we may consider any chain merely as a submatrix of $\boldsymbol{M}$. The chain is a *cycle* if each of the corresponding numbers is $\equiv 0$ (mod 2). The columns $C_{i_1}, \cdots, C_{i_p}$ are *independent* (mod 2) if there exists no set of integers $\alpha_1, \cdots, \alpha_n$ not all $\equiv 0$ (mod 2), with $\alpha_i = 0$ $(i \neq i_1, \cdots, i_p)$, such that $\Sigma\alpha_i C_i$ is a cycle, i. e. if no non-null subset of $C_{i_1}, \cdots, C_{i_p}$ is a cycle. Using this definition, the terms base, circuit, rank, nullity etc. (mod 2) can be defined as in Part I.

Let M be a set of elements $e_1, \cdots, e_n$ corresponding to $C_1, \cdots, C_n$ in $\boldsymbol{M}$, and let $e_{i_1} + \cdots + e_{i_p}$ be a circuit in M if and only if $C_{i_1}, \cdots, C_{i_p}$ is a circuit in $\boldsymbol{M}$. We shall show that *M is a matroid satisfying* (C*) *and the definitions of cycle in M and $\boldsymbol{M}$ agree.*

We show first that each circuit is a cycle in $\boldsymbol{M}$. If $C_{i_1}, \cdots, C_{i_p}$ is a circuit, then these columns are dependent; hence $\Sigma\alpha_i C_i$ is a cycle, with $\alpha_i = 0$ $(i \neq i_1, \cdots, i_p)$. Moreover $\alpha_i = 1$ $(i = i_1, \cdots, i_p)$, for otherwise a proper subset of $C_{i_1}, \cdots, C_{i_p}$ would be dependent. Hence $C_{i_1} + \cdots + C_{i_p}$ is a cycle. Next, any sum (mod 2) of circuits is a cycle, evidently. Next we prove (C*). Suppose $Q = C_{i_1} + \cdots + C_{i_p}$ is a cycle. Let $(k_1, \cdots, k_q)$ be a minimal subset of $(i_1, \cdots, i_p)$ such that $P = C_{k_1} + \cdots + C_{k_q}$ is a cycle; then P is a circuit. $Q - P$ is a cycle; from it we extract a circuit, just as above, etc. It follows from (C*) that the definitions of cycle in M and $\boldsymbol{M}$ agree. Theorems 33, 34 and 35 now apply to $\boldsymbol{M}$ also.

We are now ready to prove the final theorem:

THEOREM 37. *Let M be any matroid satisfying* (C*). *Suppose $\rho(M) = n$, and $e_1 + \cdots + e_{n-q}$ is a base. Then if $\boldsymbol{M}_1$ is any matrix of integers* (*mod* 2) *with $n - q$ columns which are independent* (*mod* 2), *columns $C_{n-q+1}, \cdots, C_n$ can be adjoined in a unique manner to $\boldsymbol{M}_1$, forming a matrix $\boldsymbol{M}$ of which the corresponding matroid is M.*

Let $P_1, \cdots, P_q$ be a strict fundamental set of circuits in M with respect to $e_{n-q+1}, \cdots, e_n$ (Theorem 9). Say $P_1 = e_{i_1} + \cdots + e_{i_p} + e_{n-q+1}$. Set $C_{n-q+1} \equiv C_{i_1} + \cdots + C_{i_p}$ (mod 2); this determines C_{n-q+1} as a column of 0's and 1's so that $P'_1 = C_{i_1} + \cdots + C_{i_p} + C_{n-q+1}$ is a circuit. (P'_1 is a cycle; (C*) shows that it is a single circuit, as $C_1 + \cdots + C_{n-q}$ contains no circuit.) C_{n-q+1} evidently must be chosen in this manner. We choose the remaining columns of $\boldsymbol{M}$ similarly. Let M' be the matroid corresponding to $\boldsymbol{M}$. Then $P'_1, \cdots, P'_q$ is a strict set of circuits in M'. These same sets form a strict set in M; hence, by Theorem 35, the circuits in M' correspond to those in M. Consequently $M' = M$, completing the proof.

We end by noting that the matroid M' of § 16 satisfies Postulate (C*) but corresponds to no linear graph. For letting 123 be a base and (16.2) a fundamental set of circuits and determining the matroid as in Theorem 36, we come out with exactly M'. A corresponding matrix of integers mod 2 is constructed from (16.3) with $a = b = c = d = 1$; we interchange rows and columns in the left-hand portion, leave out the last row and column of the right-hand portion, and interchange these two parts. (The relation $2a = 0$ is of course true mod 2.)

On the other hand, it is easily seen that if the element 7 is left out, there is a corresponding graph, which must be of the following sort: It has four vertices a, b, c, d, and the arcs corresponding to the elements $1, \cdots, 6$ are

$$ab, \quad ac, \quad ad, \quad bc, \quad bd, \quad cd.$$

There is no way of adding the required seventh arc.

The problem of characterizing linear graphs from this point of view is the same as that of characterizing matroids which correspond to matrices (mod 2) with exactly two ones in each column.

HARVARD UNIVERSITY.

ON THE ASYMPTOTIC DISTRIBUTION OF THE REMAINDER TERM OF THE PRIME-NUMBER THEOREM.

By Aurel Wintner.

The result of the present note is to the effect that the Riemann hypothesis is equivalent not only with the best possible order of the remainder term of the prime-number theorem but—on a proper scale—also with a generalized almost-periodic behavior of this remainder term. This means that the *formally trigonometrical* development of the remainder term is a *Fourier* development, which implies, in particular, the existence of an asymptotic distribution function.

Let $\rho_1, \rho_2, \cdots$ be the sequence of distinct zeros of $\zeta(s)$ in the upper half-plane, so that

$$\rho_k = 1/2 + i\gamma_k, \quad \gamma_{k+1} > \gamma_k > 0 \qquad (k = 1, 2, \cdots) \tag{1}$$

by assumption. Let n_k denote the multiplicity of the zero ρ_k and let

$$\begin{aligned} &\phi_0(x) \equiv 0, \quad \phi_m(x) = x^{-1/2}(A_m + \bar{A}_m), \text{ where } x > 1, \\ &A_m = \sum_{k=1}^{m} n_k x^{\rho_k}/\rho_k, \qquad (m = 1, 2, \cdots). \end{aligned} \tag{2}$$

It is known [1] that

$$\phi(x) = \lim_{m=\infty} \phi_m(x) \tag{3}$$

exists and that on placing, as usual,

$$\psi(x) = \sum_{p^l \leq x} \log p$$

the "explicit formula" of the prime-number theory may be written as

$$x - \psi(x) = x^{1/2}\phi(x) + \log[2\pi(1 - x^{-2})^{1/2}], \tag{4}$$

where $x \neq p^n$; at the discontinuity points, $x = p^n$, of $\psi(x)$ one has to replace $\psi(x)$ by the arithmetical mean of $\psi(x+0)$ and $\psi(x-0)$. The rôle of (1) for the distribution of the prime numbers is [2] that of implying for the remainder term (4) of the prime-number theorem $\psi(x) \sim x$ the appraisal $x^{1/2}O(x^{\epsilon})$ for any $\epsilon > 0$, and even the appraisal

[1] Cf. E. Landau, *Vorlesungen über Zahlentheorie*, Leipzig, 1927, Theorem 452.

[2] *Ibid.*, Theorem 453.

(5) $$\phi(x) = O(\log^2 x).$$

Finally, (1) is equivalent [3] also with

(6) $$(\log \omega)^{-1} \int_2^\omega (1 - \psi(x)/x)^2 \, dx \to \sum_{k=1}^\infty 2n_k^2/|\rho_k|^2, \quad \omega \to \infty,$$

a relation which may be written, according to (4), in the form

(6 a) $$(\log \omega)^{-1} \int_2^\omega x^{-2}\{x^{1/2}\phi(x) + \log[2\pi(1 - x^{-2})^{1/2}]\}^2 \, dx \\ \to 2 \sum_{k=1}^\infty n_k^2/|\rho_k|^2, \quad \omega \to \infty.$$

Now not only (6 a) holds but also

(7) $$(\log \omega)^{-1} \int_2^\omega x^{-2}\{x^{1/2}[\phi(x) - \phi_{m-1}(x)] + \log[2\pi(1 - x^{-2})^{1/2}]\}^2 \, dx \\ \to 2 \sum_{k=m}^\infty n_k^2/|\rho_k|^2; \quad (m = 1, 2, \cdots) \quad (\omega \to \infty).$$

If $m = 1$, then (7) reduces to (6 a) or (6) in virtue of (2). While for $m \neq 1$ the relation (7) is not clear from (6 a), (3) and (2), a glance at the proof of (6) shows that the proof of (7) needs but a repetition of the proof of (6 a), so that the proof of (7) will be omitted.

On denoting the integrand of (7) by

(7 a) $$x^{-1}\{\phi(x) - \phi_{m-1}(x)\}^2 + D_m(x),$$

it follows from (1), (2) and (5) that

$$D_m(x) = 2x^{-3/2}[\phi(x) - \phi_{m-1}(x)]\log[2\pi(1 - x^{-2})^{1/2}] + x^{-2}\log^2[2\pi(1 - x^{-2})^{1/2}] \\ = 2x^{-3/2}[O(\log^2 x) + O(1)]O(1) + x^{-2}O(1)^2 = O(x^{-5/4});$$

hence

$$\int_2^\omega D_m(x)\,dx = \int_2^\omega O(x^{-5/4})\,dx = O(1) = o(\log \omega).$$

Thus it is clear from (7), (7 a) that, for every fixed m,

$$(\log \omega)^{-1} \int_2^\omega x^{-1}[\phi(x) - \phi_{m-1}(x)]^2 \, dx \to 2 \sum_{k=m}^\infty n_k^2/|\rho_k|^2, \quad \omega \to \infty,$$

i. e.,

$$T^{-1} \int_1^T [\phi(e^x) - \phi_{m-1}(e^x)]^2 \, dx \to 2 \sum_{k=m}^\infty n_k^2/|\rho_k|^2, \quad T \to \infty.$$

---

[3] *Ibid.*, Theorems 476 and 477. This result is due to Cramér.

Hence, on placing

$$M\{g\} = \lim_{T=\infty} \int_1^T g(x)\,dx/T, \tag{8}$$

one has

$$M\{(f - s_{m-1})^2\} = 2 \sum_{k=m}^{\infty} n_k^2/|\rho_k|^2, \tag{9}$$

where
$$s_m(x) = \phi_m(e^x), \quad f(x) = \phi(e^x),$$
so that

$$s_m(x) = \sum_{k=1}^{m} n_k(e^{i\gamma_k x}/\rho_k + e^{-i\gamma_k x}/\bar{\rho}_k) \tag{10}$$

and

$$f(x) = \sum_{k=1}^{\infty} n_k(e^{i\gamma_k x}/\rho_k + e^{-i\gamma_k x}/\bar{\rho}_k) \tag{11}$$

in virtue of (1), (2) and (3). The relations (4) and (5) take the form

$$e^x - \psi(e^x) = e^{x/2} f(x) + \log 2\pi + O(e^{-2x}), \quad f(x) = O(x^2). \tag{11 a}$$

Now, from (9),

$$M\{(f - s_m)^2\} \to 0, \quad m \to \infty. \tag{12}$$

Due to (10) and (11), the relation (12) might be expressed by saying that the trigonometrical series (11) not only is convergent but that it is the Fourier series of the function which it represents, i. e., that

$$f(x) \sim \sum_{k=1}^{\infty} n_k(e^{i\gamma_k x}/\rho_k + e^{-i\gamma_k x}/\bar{\rho}_k), \tag{12 a}$$

the equivalence sign $\sim$ being understood in the Besicovitch sense.[4] It must, however, be mentioned that the averaging process (8) operates not in the symmetric range $[-T, T]$ but only in the upper half of it; to the lower half of this range there corresponds the range $0 < x < 1$ of $\psi(x) \equiv 0$, where the behavior of the series $\phi(x)$ is just as intricate as in the range $1 < x < \infty$.

Since, however, (11) is a pure sine series, the formal difficulty just mentioned may be avoided by *defining* the function $f(x)$ for $-\infty < x < -1$ by $f(x) = -f(-x)$ and for $-1 \leqq x \leqq 1$ arbitrarily. Then $T^{-1}\int_1^T$ in (8) may be replaced by $(2T)^{-1}\int_{-T}^{T}$, so that the function $f(x)$ occurring in the fundamental formula (11 a) belongs to the Besicovitch class B^2 in virtue of (12). As a consequence of this fact, it follows from the Besicovitch theory

[4] A. S. Besicovitch, *Almost periodic functions*, Cambridge, 1932, Chap. II.

that *the coefficients*[5] n_k/ρ_k, $n_k/\bar{\rho}_k$ *of the series* (11) *may be represented in the Fourier-Bohr manner as averages.* It is clear from the definition of $f(x)$ for $x < -1$ that these expressions of the coefficients hold also if $M\{\ \}$ is understood in the sense (8). Unfortunately, nothing is known about the diophantine nature[6] of the frequencies $\pm\gamma_n$ of the Fourier expansion (12 a); in particular, it is not known in what manner the frequencies are generated by a basis of linearly independent numbers.

If $g(x)$ is a real-valued measurable function defined for $1 < x < \infty$, and if, for a given $T > 1$ and a given real number ξ, one denotes by $[g(x) < \xi; T]$ the set of those points x of the interval $1 < x < T$ at which $g(x) < \xi$, the function $g(x)$ is said to possess an asymptotic distribution function $\sigma = \sigma(\xi)$ if at every continuity point ξ of this σ the relation

$$T^{-1} \operatorname{meas} [g(x) < \xi; T] \to \sigma(\xi), \quad (T \to \infty)$$

holds, and

$$\sigma(-\infty) = 0, \quad \sigma(+\infty) = 1.$$

Thus $\sigma(\xi)$ is monotone and not everywhere constant. The limit defining $\sigma(\xi)$ is zero for every ξ if, as x increases indefinitely, $g(x)$ is very often very large; hence in such a case there does not exist an asymptotic distribution function. Now $s_m(x)$ is, according to (10), real-valued and almost-periodic in the Bohr sense and has therefore[7] an asymptotic distribution function. Hence it follows[8] from (12) that $f(x)$ also possesses an asymptotic distribution function. This fact expresses a certain amount of regularity in the fluctuations of $f(x)$ and implies, in particular, that $|f(x)|$ cannot be very often very large. On the other hand,[9]

$$\text{(11 b)} \qquad f(x) = \Omega_{\pm}(\log\log x), \quad x \to \infty,$$

so that neither $f(x)$ nor $-f(x)$ is less than a positive constant. While the

---

[5] It is not known if $n_k = 1$ for every k.

[6] For a property of the numbers γ_n which is, however, not of an arithmetical nature, cf. pp. 101-102 of this volume.

[7] A. Wintner, "Diophantische Approximationen und Hermitische Matrizen," I, *Mathematische Zeitschrift*, vol. 30 (1929), pp. 310-312. In the following year, Jessen, and Bohr and Jessen, also proved the existence of asymptotic distribution functions. Cf. also the programmatic address of Bohr in the *Proceedings of the 5th Skandinavian Congress* (1922). For the recent development of the distribution theory, cf. B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, July, 1935.

[8] Cf. B. Jessen and A. Wintner, *loc. cit.*, Theorem 24.

[9] Cf. E. Landau, *op. cit.*, Theorem 472, or H. Bohr, *loc. cit.* (11 b) is due to Littlewood.

asymptotic distribution function cannot be everywhere constant, it seems to be difficult to decide whether it is nowhere constant[10] or whether $f(x)$ "dislikes" some regions $a < f(x) < b$. The answer to this question might depend[11] on the one mentioned at the end of the previous paragraph.

The relation (12 a) may be expressed also in terms of the Dirichlet series

$$F(s) = \sum_{k=1}^{\infty} n_k/\gamma_k e^{-\gamma_k s}, \quad \Re s > 0.$$

It is known[12] that (11 b) depends on the behavior of

$$\Im F(s) \text{ as } \Re s \to +0.$$

Now $\Im F(s)$ on the boundary line $\Re s = 0$ not only is a convergent trigonometrical series but is also the Fourier series of the function which it represents. In fact, the number of zeros of $\zeta(s)$ in the strip $0 < \Im s < T$ is $O(T \log T)$, even if one counts every zero according to its multiplicity. This implies the convergence of the series

$$\sum_{k=1}^{\infty} n_k^2/|\gamma_k \rho_k|.$$

Hence it follows from (11) that the series representing the sum of $f(x)$ and $2\Im F(ix)$ is a trigonometrical series which possesses a convergent majorant uniformly for all x and is therefore almost-periodic in the sense of Bohr. Accordingly the trigonometrical series

$$-\sum_{k=1}^{\infty} n_k/\gamma_k \sin \gamma_k x,$$

which defines the function $\Im F(ix)$, belongs to this function as its Fourier series (B^2). In particular, $\Im F(ix)$ is a function of class B^2 and has therefore an asymptotic distribution function.

THE JOHNS HOPKINS UNIVERSITY.

---

[10] Cf. A. Wintner, "Remarks on the Ergodic Theorem of Birkhoff," *Proceedings of the National Academy of Sciences*, vol. 18 (1932), p. 251.

[11] Cf., in this connection, B. Jessen and A. Wintner, *loc. cit.*

[12] Cf. E. Landau, *op. cit.*, Theorem 470.

ON THE EXACT VALUE OF THE BOUND FOR THE REGULARITY OF SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS.

By Aurel Wintner.

Let $f(z, w)$ be regular-analytic and bounded in the four-dimensional domain $|z| < a$, $|w| < b$, and let $w = w(z)$ be the solution of $dw/dz = f(z, w)$ which vanishes at $z = 0$. Let M denote the least upper bound of $|f(z, w)|$ in the domain $|z| < a$, $|w| < b$. It is known that there exists a bound $\Gamma = \Gamma(a, b, M)$ which is independent of the particular choice of $f(z, w)$ and is such that $w(z)$ is regular-analytic in the circle $|z| < \Gamma$. In fact, the method of successive approximations yields the estimate

$$\Gamma(a, b, M) \geqq \min(a, b/M). \tag{1}$$

The necessity of the limitation $|z| < a$ is obvious from the case where $f(z, w)$ is independent of w and has singularities on the circle $|z| = a$. On the other hand, the necessity of the limitation $|z| < b/M$ is not evident. In fact, the latter limitation is introduced into the proof of (1) only for a somewhat artificial reason,—in order to assure the possibility of successive substitutions into f.

It turns out, however, that the trivial, and *a priori* artificial, appraisal (1) cannot be improved, i. e., that *the value of the best bound* $\Gamma(a, b, M)$ *is precisely* $\min(a, b/M)$. This situation seems to be unexpected insofar as efforts have been made [1] to improve the lower estimate (1) of $\Gamma(a, b, M)$. In reality, these efforts succeeded only by imposing additional restrictions on $f(z, w)$. Such a restriction is that $f(z, w)$ satisfies a uniform Lipschitz condition in the open domain $|z| < a$, $|w| < b$; and the corresponding improved estimate of the regularity radius of $w(z)$ depends * not only on a, b, M but also on the Lipschitz constant. A proof of the upper estimate

$$\Gamma(a, b, M) \leqq \min(a, b/M), \tag{2}$$

which clears up this situation, runs as follows.

If $a \leqq b/M$ then (2) is obvious from the case where $f(z, w)$ is independent of w. In order to prove that (2) holds also when $a > b/M$, it is clearly sufficient to show that for any *given* pair of numbers b, M and for any

[1] Cf. P. Painlevé, *Encyklopaedie der Mathematischen Wissenschaften*, vol. 1, Part I_2, p. 194 and p. 200.

given number r, where $r > b/M$, there exists a function f which is independent of z and possesses the following properties:

(i) $f(w)$ is regular-analytic and bounded in the circle $|w| < b$;

(ii) the least upper bound of $|f(w)|$ in this circle is M;

(iii) the function $w = w(z)$ for which $dw/dz = f(w)$ and $w(0) = 0$ has in the circle $|z| < r$ a singularity.

Now the function

$$f(w) = M[(1 + w/b)/2]^{1/n} \tag{3}$$

satisfies all these conditions if n is sufficiently large, larger than a number depending on r. In fact, the solution $w(z)$ belonging to (3) is

$$w(z) = b[(1 + z/C_n)^{n/(n-1)} - 1]$$

where

$$C_n = (1 - 1/n)^{-1} 2^{1/n} b/M.$$

Thus $z = -C_n$ is a singular point of $w(z)$ and tends to $z = -b/M$ when $n \to +\infty$. This proves (iii), while (i) and (ii) are satisfied by (3) for any n.

The Johns Hopkins University.

ON SYMMETRIC BERNOULLI CONVOLUTIONS.

By RICHARD KERSHNER and AUREL WINTNER.

A class of symmetric Bernoulli convolutions which are regular analytic in the whole plane or in a strip containing the real axis or which possess, at least, a high degree of smoothness along the real axis, has recently been considered by one of the authors.[1] The present note deals mainly with the other extreme case, where the convolution possesses but a very low degree of smoothness. One class of convolutions which will be considered includes, for instance, the well known Cantor function; and other functions which have been treated in the literature also occur. These and some other examples have been collected in a joint paper of B. Jessen and one of the present authors.[2] The present note attempts a systematic treatment of a type of these symmetric convolutions with a low degree of smoothness. Apart from the theory of infinite convolutions, the functions to be considered are of interest from the point of view of the theory of real functions. The dominating feature of some of the convolutions in question is the homogeneous character of their spectra and a corresponding homogeneity of the mapping involved. In particular, one is lead to absolutely continuous convolutions which might be termed length-preserving with respect to a nowhere dense set of positive measure. It turns out that Bernoulli convolutions of this type are identical with the functions $\phi(x)$ considered by Hausdorff[3] in connection with his fractional dimension theory. While Hausdorff is mainly interested in the case where the Lebesgue measure, which will be denoted by $\mu(E)$, is zero, his results hold for the case of a positive Lebesgue measure also, a case with which the present paper is mainly concerned. A class of Bernoulli convolutions which might be termed complementary to the case of the Hausdorff functions $\phi(x)$ also is considered.

Let $\beta(x)$ denote the symmetric Bernoulli distribution of standard de-

---

[1] A. Wintner, "On analytic convolutions of Bernoulli distributions," *American Journal of Mathematics*, vol. 56 (1934), pp. 659-663; "On symmetric Bernoulli convolutions," *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 137-138.

[2] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta-function," *Transactions of the American Mathematical Society*, vol. 37 (1935), § 6, Theorem 11.

[3] F. Hausdorff, "Dimension und äusseres Mass," *Mathematische Annalen*, vol. 79 (1919), pp. 157-179.

viation 1 so that $\beta(x)$ is 0, 1/2 or 1 according as x is on the left, in the interior or on the right of the interval $-1 < x < +1$. Thus $\beta(x/a)$, where $a > 0$, also is a symmetric Bernoulli distribution function; the distribution function

$$\tfrac{1}{2}(1 + \operatorname{sign} x), \tag{1}$$

which belongs to $a = +0$, will not be considered as a Bernoulli distribution function. The infinite Bernoulli convolution

$$\sigma(x) = \beta(x/a_1) * \beta(x/a_2) * \cdots \tag{2}$$

is convergent if and only if

$$\sum_{n=1}^{\infty} a_n^2 < +\infty, \tag{3}$$

and the convergence of (2) implies its absolute convergence.[4] It will always be supposed that (3) is satisfied. The function (2) always is continuous.[5] Further,[6] if $\sigma(x)$ is not absolutely continuous, it is singular, i. e. such that $\sigma'(x) = 0$ almost everywhere. The spectrum S of $\sigma(x)$, defined as the set of points x in the vicinity of which σ is not constant, consists[7] of those points x which are representable in the form

$$\sum_{n=1}^{\infty} \pm a_n, \tag{4}$$

where the signs depend on n in an arbitrary way, the only restriction being that the series be convergent. Hence S is a bounded set or the whole real axis according as the condition

$$\sum_{n=1}^{\infty} a_n < +\infty \tag{5}$$

is or is not satisfied. Examples show[8] that σ may be singular or absolutely continuous whether (5) is satisfied or not, so that all four possibilities actually occur. The set S is always perfect, since the set of points in the vicinity of which a continuous function is not constant is either perfect or empty. On denoting by ρ_n the infinite convolution

$$\rho_n(x) = \beta(x/a_{n+1}) * \beta(x/a_{n+2}) * \cdots, \tag{6}$$

so that ρ_n tends, as $n \to +\infty$, to the distribution function (1), either all functions

---

[4] B. Jessen and A. Wintner, *loc. cit.* [5] *Ibid.* [6] *Ibid.* [7] *Ibid.*

[8] *Ibid.*, examples 1, 3, 5, and 6.

$$(7) \qquad \rho_0 = \sigma, \rho_1, \rho_2, \cdots$$

are singular or all are absolutely continuous. For if $\rho_n(x)$ is singular, then the derivative of

$$\rho_{n-1}(x) = \rho_n(x) * \beta(x/a_n) = \tfrac{1}{2}[\rho_n(x+a_n) + \rho_n(x-a_n)]$$

is zero almost everywhere; and if $\rho_n(x)$ is absolutely continuous, then so is $\rho_{n-1}(x)$, since absolute continuity cannot be lost by the convolution process.[9] It may be mentioned that if S_n denotes the spectrum of ρ_n, then either all sets

$$(8) \qquad S_0 = S, S_1, S_2, \cdots$$

are nowhere dense or none are nowhere dense. For if S_n contains an interval, then so does each of the sets[10] $S_n - a_n$ and $S_n + a_n$, the logical sum of which is S_{n-1} in virtue of (4); further, if S_n contains an interval, then so does S_{n+1}. For if I be an interval in S_n, then $I - a_{n+1}$ either has an interval in common with the perfect set S_{n+1} or contains a subinterval $J - a_{n+1}$ which does not contain any point of S_{n+1}. In the latter case $J + a_{n+1}$ is contained in S_{n+1} by the definition of S_n in terms of S_{n+1}, so that in either case S_{n+1} contains an interval.

From now on it will be supposed that S is bounded, i. e., that (5) is satisfied, so that one may introduce the remainders

$$(9) \qquad r_n = \sum_{m=n+1}^{\infty} a_m \qquad (n = 0, 1, 2, \cdots).$$

The following theorem will now be proven:

If

$$(10) \qquad a_n > \sum_{m=n+1}^{\infty} a_m = r_n$$

---

[9] In fact, absolute continuity, in the case of a distribution function $\psi(x)$, means that there exists for every $\epsilon > 0$ a $\delta = \delta(\epsilon)$ such that

$$\sum_{k=1}^{\infty} |\psi(x'_k) - \psi(x''_k)| < \epsilon \quad \text{whenever} \quad \sum_{k=1}^{\infty} |x'_k - x''_k| < \delta.$$

Since the latter inequality implies that $\sum_{k=1}^{\infty} |(x'_k - y) - (x''_k - y)| < \delta$ for any y, it also implies that $\sum_{k=1}^{\infty} |\psi(x'_k - y) - \psi(x''_k - y)| < \epsilon$ for any y. Hence it implies that

$$\sum_{k=1}^{\infty} \left| \int_{-\infty}^{+\infty} \psi(x'_k - y)\, d\omega(y) - \int_{-\infty}^{+\infty} \psi(x''_k - y)\, d\omega(y) \right| \leqq \epsilon \int_{-\infty}^{+\infty} d\omega(y) = \epsilon$$

for any distribution function ω or that $\psi * \omega$ is absolutely continuous for any distribution function ω.

[10] By $E + c$ is meant the set of points representable in the form $x + c$, where x is a number contained in E.

for every n, then the spectrum S of the infinite Bernoulli convolution (2) is nowhere dense and has the measure

$$(11) \qquad \mu(S) = 2\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} 2^n (a_n - \sum_{m=n+1}^{\infty} a_m) = 2 \lim_{n=\infty} 2^n r_n,$$

which may or may not be zero. If $\mu(S) > 0$ and X denotes the interval $[-\infty, x]$, then

$$(12) \qquad \sigma(x) = \mu(SX)/\mu(S),$$

i. e., the measure of the σ-image of an x-interval is proportional to the portion of the nowhere dense spectrum contained in the interval. In particular, $\sigma(x)$ is singular or absolutely continuous according as $\mu(S) = 0$ or $\mu(S) > 0$. If it is absolutely continuous, then its density is bounded [11] since σ satisfies a uniform Lipschitz condition.

The fact that $\mu(S) > 0$ implies absolute continuity is of interest since if (10) is not satisfied, then $\mu(S) > 0$ is necessary but not sufficient for the absolute continuity of $\sigma(x)$; cf. (23).[11a] That in the case (10) the condition $\mu(S) > 0$ is sufficient for absolute continuity is clear from the relation (12), since (12) implies the uniform Lipschitz condition

$$| \sigma(x') - \sigma(x'') | \leqq C \, | x' - x'' |,$$

the best value of C being $1/\mu(S)$. The second representation of $\mu(S)$ given in (11) shows that $\mu(S) = 0$ or $\mu(S) > 0$ according as $r_n = o(2^{-n})$ is or is not satisfied, and that (10) implies the existence of $\lim 2^n r_n$.

It is seen from (9) that (10) may be written in the form

$$(13) \qquad r_n > 2r_{n+1} \qquad (n = 0, 1, 2, \cdots).$$

This clearly implies the possibility of a successive construction of open subintervals J of the closed interval $[-r_0, r_0]$ as follows. Let J_{12} denote the open interval which is symmetric with respect to the mid-point of $[-r_0, r_0]$ and is of length $2(r_0 - 2r_1)$. From each of the two closed intervals K_1, K_2 which constitute $[-r_0, r_0] - J_{12}$ one may remove an open interval of length $2(r_1 - 2r_2)$ and having the mid-point of K_i, where $i = 1, 2$, as mid-point. This is possible since K_1 and K_2 are each of length $4r_1 > 2(r_1 - 2r_2)$. Let J_{14} and J_{34} denote the open intervals thus removed and let J_{14} be the one

[11] Up to a set of measure zero.

[11a] For another example of this type cf. A. Denjoy, "Sur quelques points de la théorie des fonctions," *Comptes Rendus*, vol. 194 (1932), pp. 44-46, and H. Minkowski, *Gesammelte Abhandlungen*, vol. 2 (1911), pp. 50-51 and fig. 7.

which is to the left of J_{12}. From each of the four closed intervals which constitute

$$[-r_0, r_0] - J_{12} - J_{14} - J_{34}$$

one may remove open intervals $J_{18}, J_{38}, J_{58}, J_{78}$ in the same symmetric manner as J_{14} and J_{34} have been removed from K_1 and K_2, it being understood that each of the four intervals J_{k8} is of length $2(r_2 - 2r_3)$ and that J_{k8} is on the left of J_{h8} if $k < h$. On continuing this process one obtains for every n

$$(14)\quad \begin{array}{l} 2^n \text{ intervals } J_{k2^{n+1}} \text{ of length } 2(r_n - 2r_{n+1}), \\ \text{where } k = 1, 3, 5, \cdots, 2^{n+1} - 1 \end{array} \qquad (n = 0, 1, 2, \cdots).$$

It is convenient to write the double subscript of J_{pq} as a fraction by placing $J_{pq} = J_{p/q}$ so that J_t is defined for every number t of the form

$$(15)\qquad t = \sum_{j=1}^{m} b_j/2^j, \text{ where } b_j = 0 \text{ or } b_j = 1,$$

i. e., for every number of the interval $0 < t < 1$ having a finite dyadic development. J_t and J_u are, by their successive construction, disjoint if $t \neq u$. Now it is easy to verify [12] from the definition of $\beta(x/a_n)$ and from that of the convolution operator $*$ that

$$(16)\qquad \sigma(x) \equiv t \text{ if } x \text{ is in } J_t.$$

Since the points (15) lie dense in the interval $0 < t < 1$ and since the distribution function $\sigma(x)$ is everywhere continuous, it follows that every subinterval of $[-r_0, r_0]$ contains a J_t. Consequently, the set

$$(17)\qquad [-r_0, r_0] - \sum_t J_t,$$

where t runs through all values (15), is nowhere dense and consists of the cluster points of the endpoints of the open intervals J_t. Hence it is clear from (16) and from the definition of the spectrum S that the set (17) is contained in S. Since S is a subset of $[-r_0, r_0]$ in virtue of (4), the set (17) is precisely S. Accordingly, J_t and J_u being disjoint if $t \neq u$,

$$\mu(S) = \mu([-r_0, r_0]) - \mu(\sum_t J_t) = 2r_0 - \sum_t \mu_t(J_t),$$

so that

$$\mu(S) = 2r_0 - \sum_{n=0}^{\infty} 2^n (r_n - 2r_{n+1})$$

in virtue of (14). On comparing this with (9) one obtains (11).

---

[12] Cf. B. Jessen and A. Wintner, *loc. cit.*

It must now be shown that in the case $\mu(S) > 0$ the relation (12) holds.[13] Since $\sigma(x)$ is non-decreasing and continuous, it is sufficient to verify (12) for a dense set of points x. Now $[-r_0, r_0]$ contains S, so that (12) is trivial if x is not in $[-r_0, r_0]$. Since $\sum_t J_t$ is dense in $[-r_0, r_0]$, it follows that it is sufficient to verify (12) for the points x of a J_t. Let x be in J_t and let $t = k/2^m$. Then the $2^m - 1$ intervals J_t, where $t = j/2^n$ and $j = 1, 3, 5, \cdots, 2^{n-1}$; $n = 1, 2, 3, \cdots, m$, decompose S into 2^m congruent parts, each of which has the measure $2^{-m}\mu(S)$ since the intervals J_t have been removed symmetrically. Since there are, among the 2^m congruent parts of measure $2^{-m}\mu(S)$, exactly k on the left of the point x, one has

$$\mu(SX) = k2^{-m}\mu(S).$$

This proves (12) since $t = k/2^m$, and $\sigma(x) = t$ by (16).

The Hausdorff theory[14] of λ-measure and its further development by Besicovitch[15] allow, of course, an analysis of the case $\mu(S) = 0$ also.

As an illustration of the theorem, let

$$(18) \qquad a_n = Aa^n + Bb^n, \text{ where } 0 < a < b < 1,\ A > 0,\ B \geqq 0.$$

It is easily verified that (10) is satisfied if and only if

$$(19) \qquad b \leqq 1/2$$

and that (11) gives $\mu(S) = 0$ or $\mu(S) = 2B$, according as $b < 1/2$ or $b = 1/2$. Thus if

$$(20) \qquad a_n = B(1/2)^n + Aa^n, \text{ where } 0 < a < 1/2,\ B > 0,\ A > 0,$$

then (2) is absolutely continuous with a nowhere dense set of positive measure as spectrum and is represented by the formula (12).

The infinite convolution (2) belonging to the sequence (18) in the case $A = 1$, $B = 0$ will be denotel by $\sigma_a(x)$, so that

$$(21) \qquad \sigma_a(x) = \beta(x/a) * \beta(x/a^2) * \beta(x/a^3) * \cdots \quad (0 < a < 1).$$

Since (19) takes, in the case $B = 0$, the form $a < 1/2$ in virtue of $a < b$, the function $\sigma_a(x)$ is singular with a spectrum S of zero measure if $a < 1/2$. In particular, $\sigma_{1/3}(x)$ is the usual Cantor function considered by Lebesgue.

[13] Cf. F. Hausdorff, *loc. cit.*, § 11.

[14] F. Hausdorff, *loc. cit.*, §§ 10-12.

[15] A. S. Besicovitch, "On linear sets of points of fractional dimension," *Mathematische Annalen*, vol. 101 (1929), pp. 161-193.

On the other hand, $\sigma_{1/2}(x)$ is the distribution function of the "Abrundungsfehler," i. e.,

$$\sigma_{1/2}(x) = (x+1)/2 \text{ if } -1 \leqq x \leqq 1$$

and $S = [-1, 1]$, so that $\sigma_{1/2}(x)$ is absolutely continuous with a bounded density. This example shows that on replacing $>$ in (10) by $\geqq$, the spectrum S of (2) may become an interval. Since the sequence $\{(1/2)^n\}$ consists of the two sequences $\{(1/4)^n\}$ and $\{(1/4)^n/2\}$, it is clear that

$$\sigma_{1/4}(x) * \sigma_{1/4}(2x) = \sigma_{1/2}(x); \tag{22 a}$$

this relation is an instance of the fact that the convolution of two singular Bernoulli convolutions may be absolutely continuous.[16] On the other hand,

$$\sigma_{1/4}(x) * \sigma_{1/4}(x) \tag{22 b}$$

is singular. This is shown in the same way [17] as for $\sigma_{1/3}(x) * \sigma_{1/3}(x)$. The interest of the latter example lies in the fact that although it is singular, the spectrum is an interval, as will follow from (23). Besides, (23) will show that the spectrum of $\sigma_a(x)$ is an interval not only in the limiting case $a = 1/2$ but in the case $1/2 < a < 1$ as well.

If $>$ in (10) be replaced by $\leqq$, then S becomes connected:

$$S = [-r_0, r_0] \text{ if } a_n \leqq r_n \qquad (n = 1, 2, \cdots). \tag{23}$$

This is easily seen from (4) and (9) by an obvious extension of the proof of Riemann's theorem according to which

$$0 < a_n \to 0 \quad \text{and} \quad \sum_{n=1}^{\infty} a_n = +\infty$$

imply that every real number is representable in the form (4). The examples

$$a_n = (1/2)^n \text{ and } a_{2n} = a_{2n+1} = (1/3)^n$$

show that in the case (23) both absolutely continuous and singular convolutions (2) are possible.

If (10), i. e. $a_n > r_n$, is satisfied not for every n but only for sufficiently large values of n, then S is still nowhere dense since S_n in (8) is then nowhere dense if n is sufficiently large. If $a_n \leqq r_n$ is satisfied not for every n but only for sufficiently large n, then S_n is, according to (23), an interval if n is

---

[16] Cf. P. Lévy, "Sur les séries dont les termes sont des variables éventuelles indépendantes," *Studia Mathematica*, vol. 3 (1931), p. 153.

[17] Cf. B. Jessen and A. Wintner, *loc. cit.*, example 2.

sufficiently large, so that $S = S_0$ consists of a finite number ($\geqq 1$) of intervals in virtue of (4). The following example shows that S may consist of an arbitrarily large number of disjoint intervals.

For a given $\alpha > 1$, let $\sigma^\alpha(x)$ denote the Bernoulli convolution belonging to $a_n = n^{-\alpha}$, so that

$$(24) \qquad \sigma^\alpha(x) = \beta(x/1^\alpha) * \beta(x/2^\alpha) * \beta(x/3^\alpha) * \cdots,$$

and let S^α denote the spectrum of (24). Since $a_n = n^{-\alpha}$ satisfies $a_n \leqq r_n$ for sufficiently large n if α is fixed, the spectrum S^α consists of $N = N_\alpha$ disjoint intervals. It is easy to see that, as $\alpha \rightarrow +\infty$, the number of intervals increases indefinitely while S^α shrinks to the set consisting of the pair of points $x = \pm 1$. In this connection it is interesting to mention that[18] $\sigma^\alpha(x)$ possesses, for every fixed α, derivatives of arbitrarily high order for every x, although $\sigma^\alpha(x)$ cannot, of course, be analytic at the end points of the N_α intervals which constitute S^α. It is not known whether $\sigma^\alpha(x)$ is or is not analytic in the interior of these intervals. Since every n may be written uniquely in the form $n = 2^k(2m + 1)$, it is clear from (24) that[19]

$$(25) \quad \sigma^\alpha(x) = \sigma_c(x/1^\alpha) * \sigma_c(x/3^\alpha) * \sigma_c(x/5^\alpha) * \cdots, \text{ where } c = (1/2)^\alpha < 1/2,$$

so that $\sigma_c(x)$ is singular with a spectrum of zero measure. It may be mentioned that if α is near enough to 1, then $N_\alpha = 1$, and that $S^\alpha \rightarrow [-\infty, \infty]$ as $\alpha \rightarrow 1 + 0$.

THE JOHNS HOPKINS UNIVERSITY.

[18] A. Wintner, *loc. cit.*

[19] Cf. P. Lévy, *loc. cit.*, p. 154.

ON UNIFORM CONVERGENCE.

By J. W. THEODORE SUCKAU.

Introduction. While in Classical Analysis it is always explicitly presupposed that a convergent sequence of functions is uniformly convergent, it was first observed by Egoroff [1] that a very strong type of approximate uniform convergence is automatically present in every convergent sequence of measurable functions.

The theorem of Egoroff has led to many investigations, in particular by F. Riesz [2] who has applied the theorem in a very interesting manner to the Lebesgue theory.[3]

It is the purpose of this paper to investigate the phenomenon of uniform convergence in a general way so as to obtain a better understanding of the situation in the case of measurable functions.

I. *Uniform Convergence.* Suppose that the sequence $f_n(x)$[4] is defined and convergent on a set S. If the sequence is uniformly convergent on some subset of S, that subset is said to have the character U (and is designated by the letter U).

We introduce a class of sets, Ψ, namely the totality of subsets of S having the character U.

The class Ψ is neither the null set nor does it contain only the null set, since any subset of S with a finite number of elements is a subset of character U. Moreover, the addition of a finite number of points of S to an element of Ψ yields another element in Ψ. Hence, except in the trivial case when the sequence is uniformly convergent on S there is no largest subset in the class Ψ.

---

[1] D. Th. Egoroff, "Sur les suites des fonctions measurables," *Comptes Rendus*, Paris, vol. 152 (1910), pp. 244-246.

[2] F. Riesz, (i) "Sur l'intégrale de Lebesgue," *Acta Mathematica*, vol. 42 (1920), pp. 191-205; (ii) "Sur le théorème de M. Egoroff et sur les opérations fonctionnelles linéaires," *Acta Litt. Sci. Szeged*, vol. 1 (1922), pp. 18-25; (iii) "Elementarer Beweis des Egoroffschen Satzes," *Monatsheften für Mathematik und Physic*, vol. 25 (1928), pp. 243-248.

[3] F. Riesz, footnote [2] (i), above; *loc. cit.*, pp. 196-205.

[4] Unless otherwise stated all functions are entirely unrestricted.

In the special case when S is itself an element of Ψ it is the largest element and the class Ψ is coincident with the totality of subsets of S. However, even when this is not the case it is possible by a very simple process[5] to determine the class Ψ. This process is fundamental in the paper.

Let $S(k, \nu)$ be the subset of S on which the inequality

$$| f_n(x) - f_m(x) | \leq 1/k$$

holds for every $n, m \geq \nu$.

Take any sequence of positive integers

$$(\nu) : \nu_1, \nu_2, \nu_3, \cdots.$$

Finally, consider

$$U(\nu) = \prod_{k=1}^{\infty} S(k, \nu_k).$$

THEOREM. *The sequence $f_n(x)$ is uniformly convergent on $U(\nu)$, and conversely, if the sequence $f_n(x)$ is uniformly convergent on a subset S^* of S, then there exists a sequence (ν) such that $U(\nu)$ contains S^*. In other words the totality of sets $U(\nu)$ together with their subsets is the class Ψ.*

Proof. To prove the first part, take any $\epsilon > 0$ and pick k so that $1/k \leq \epsilon$. Now every x in $U(\nu)$ is also in $S(k, \nu_k)$, and so for every x in $U(\nu)$ it is true that $| f_n(x) - f_m(x) | \leq 1/k \leq \epsilon$ for $n, m \geq \nu_k$. Hence the sequence is uniformly convergent on $U(\nu)$.

Conversely, if $f_n(x)$ is uniformly convergent on S^* then given any positive integer k there exists a ν_k such that for all $n, m \geq \nu_k$ it is true that $| f_n(x) - f_m(x) | \leq 1/k$ for every x in S^*. Thus a sequence $(\nu) : \nu_1, \nu_2, \nu_3, \cdots$ has been defined. S^* is contained in $S(k, \nu_k)$ for every value of k and therefore S^* is contained in $\prod_{k=1}^{\infty} S(k, \nu_k) = U(\nu)$.

1.2. Now that the class Ψ has been determined it is interesting to examine its elements. We know that every finite subset of S is contained in Ψ. Under certain conditions we are assured that at least one infinite subset of S is a member of Ψ.

THEOREM. *If S is non-denumerable, then there exists a denumerably infinite subset on which the convergence is uniform.*

Proof. There is an α_1 such that $S(1, \alpha_1)$ is non-denumerable. This is

[5] Though not explicitly stated by Riesz, the process used by him is essentially the same. Footnote [2] (iii), p. 549, *loc. cit.*, p. 244.

so since $S = \sum_{\nu=1}^{\infty} S(1, \nu)$ and hence if every $S(1, \nu)$ were denumerable it would follow that S had this property, which is contrary to hypothesis.

Pick x_1 from $S(1, \alpha_1)$.

Consider the sequence now as defined only on $S(1, \alpha_1)$ which is non-denumerable. There is an α_2 such that when the construction of 1. 1 is considered with respect to $S(1, \alpha_1)$, then $S(2, \alpha_2 : 2)$ is non-denumerable. ($S(2, \alpha_2 : 2)$ is the set of points of $S(1, \alpha_1)$ where $|f_n(x) - f_m(x)| \leq \frac{1}{2}$ for $n, m \geq \alpha_2$).

Pick $x_2 \neq x_1$ from $S(2, \alpha_2 : 2)$.

Moreover, there is a β_2 such that for all $n, m \geq \beta_2$

$$|f_n(x_i) - f_m(x_i)| \leq \tfrac{1}{2} \qquad (i = 1).$$

. . . .

Consider the sequence as defined only on $S(k-1, \alpha_{k-1} : k-1)$ which is non-denumerable. There is an α_k such that when the construction of 1. 1 is considered with respect to $S(k-1, \alpha_{k-1} : k-1)$ then $S(k, \alpha_k : k)$ is non-denumerable. ($S(k, \alpha_k : k)$ is the set of points of $S(k-1, \alpha_{k-1} : k-1)$ where $|f_n(x) - f_m(x)| \leq 1/k$ for $n, m \geq \alpha_k$).

Pick $x_k \neq x_1, x_2, x_3, \cdots, x_{k-1}$ from $S(k, \alpha_k : k)$.

Moreover, there is a β_k such that for $n, m \geq \beta_k$

$$|f_n(x_i) - f_m(x_i)| \leq 1/k \qquad [i = 1, 2, 3, \cdots, (k-1)].$$

Choose $\nu_k = \overbrace{\alpha_k, \beta_k}$ [6] $: \beta_1 = \alpha_1$.

Define (ν) : $\nu_1, \nu_2, \nu_3, \cdots$.

Then $U(\nu)$ contains the set $x_1, x_2, x_3, \cdots$.

$S(k, \alpha_k : k)$ certainly contains the set $x_k, x_{k+1}, x_{k+2}, \cdots$, since $S(k, \alpha_k : k)$ contains $S(k+1, \alpha_{k+1} : k+1)$; while $x_1, x_2, x_3, \cdots, x_{k-1}$ is included in $S(k, \beta_k)$. Furthermore, $S(k, \nu_k)$ contains both the sets $S(k, \alpha_k : k)$ and $S(k, \beta_k)$ and so it contains the denumerably infinite set $x_1, x_2, x_3, \cdots$ for every value of k.

This proves the theorem.

That the non-denumerability of S in the hypothesis of the above theorem is essential is shown by the following example.

Define $$f_n(x) = \begin{cases} 0 \text{ on } 1, 1/2, 1/3, \cdots, 1/n \\ 1 \text{ on } 1/n+1, 1/n+2, 1/n+3, \cdots. \end{cases}$$

---

[6] $\overbrace{a_1, a_2, a_3, \cdots, a_n}$ is the largest of this set of values, $\underbrace{a_1, a_2, a_3, \cdots, a_n}$ the smallest.

Now $f_n(x) \to 0$ on the set $\{1/n\}$, and yet on any infinite subset the convergence is not uniform.

This example and the above theorem may be combined in the statement of the following theorem:

A necessary and sufficient condition that each convergent sequence of functions defined on S have associated with it a denumerably infinite subset of S on which the convergence is uniform, is that S be non-denumerable.

1. 3. The question arises as to whether or no this theorem is the best of its kind. We might perhaps always have a subset of character U which is non-denumerable. We are going to develop an example to show that this is not in general possible if we assume the hypothesis of the continuum, that $2^{\aleph_0} = \aleph_1$. In other words the above theorem is in a way final.

The example we shall develop is that of a sequence of functions converging to zero on the whole continuum and yet uniformly convergent on no non-denumerable subset.

Before going to the construction of the example we shall state two auxiliary theorems.

AUXILIARY THEOREM I. *If a null sequence*[7] *of functions is uniformly convergent on a set S, then there is a null sequence of numbers which dominates*[8] *the sequence of functions on S.*

This is merely a restatement of the definition of uniform convergence on S.

AUXILIARY THEOREM II. *Given any denumerable aggregate of null sequences (of numbers) there exists a null sequence dominated by none of the sequences of the aggregate.*

Proof. Suppose that the given sequences are:

$$
\begin{array}{l}
a_{11}, a_{12}, a_{13}, \cdots \to 0 \\
a_{21}, a_{22}, a_{23}, \cdots \to 0 \\
a_{31}, a_{32}, a_{33}, \cdots \to 0 \\
\cdot \quad \cdot \quad \cdot \quad \cdots \quad \cdot
\end{array}
$$

(i) Pick the first value of n, say $\nu(1)$ such that $| a_{1,\nu(1)} | < 1$. Cancel the first $\nu(1)$ columns.

[7] A null sequence is a sequence converging to zero.

[8] The sequence a_n dominates the sequence $b_n(x)$ on S if there is a subscript n_0 such that for all $n > n_0$, $a_n \geq b_n(x)$, on S.

(ii) Pick the first remaining value of n, say $\nu(2)$ such that $|a_{1,\nu(2)}| < 1/2$. Cancel the next $\nu(2) - \nu(1)$ columns.
Pick the first remaining value of n, say $\nu(3)$ such that $|a_{2,\nu(3)}| < 1/2$. Cancel the next $\nu(3) - \nu(2)$ columns.
(iii) Pick the first remaining value of n, say $\nu(4)$ such that $|a_{1,\nu(4)}| < 1/3$. Cancel the next $\nu(4) - \nu(3)$ columns.
Pick the first remaining value of n, say $\nu(5)$ such that $|a_{2,\nu(5)}| < 1/3$. Cancel the next $\nu(5) - \nu(4)$ columns.
Pick the first remaining value of n, say $\nu(6)$ such that $|a_{3,\nu(6)}| < 1/3$. Cancel the next $\nu(6) - \nu(5)$ columns.

. . . .

Now construct the sequence $b_1, b_2, b_3, \cdots$ in this wise:

$$
\begin{aligned}
&b_1 = b_2 = \cdots = b_{\nu(1)} = 1\\
&b_{\nu(1)+1} = \cdots = b_{\nu(2)} = \cdots = b_{\nu(3)} = 1/2\\
&b_{\nu(3)+1} = \cdots = b_{\nu(4)} = \cdots = b_{\nu(5)} = \cdots = b_{\nu(6)} = 1/3.\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}
$$

The sequence $\{b_n\}$ is a null sequence and it is evident that it is dominated by none of the given sequences.

We may now go to the actual *construction of the example.*

Normally order the following sets: (i) all real numbers, (ii) the totality of null sequences of numbers.

(i) $\qquad x_1, x_2, x_3, \cdots, x_\lambda, \cdots (\lambda < \Omega)$

(ii) $\qquad A_1, A_2, A_3, \cdots, A_\lambda, \cdots (\lambda < \Omega)$

where Ω is the ordinal which initiates the cardinal c.

The required sequence will be defined as an $\aleph_0$-valued function, $B(x)$.

Define $B(x_\lambda)$ to be the first sequence in (ii) which is not dominated by any A_ν for all $\nu \leq \lambda$. Suppose that the sequence is $A_{\mu(\lambda)}$.

It follows from Auxiliary Theorem II that there is a null sequence not dominated by A_ν for all $\nu \leq \lambda$; for, these null sequences constitute a denumerable aggregate, as λ is ordinally less than Ω the ordinal which initiates the cardinal c, and it is assumed that $c = 2^{\aleph_0} = \aleph_1$. Thus the definition of $B(x)$ is complete.

$B(x_\lambda)$ is possibly dominated by a null sequence, say A_ν, only for x's with subscripts $\lambda < \nu$; for if $\nu \leq \lambda$, then by definition $B(x_\lambda)$ is not dominated by A_ν. In other words $B(x)$ is possibly dominated by A_ν only for the x's in the set

$$x_1, x_2, x_3, \cdots, x_\beta, \cdots (\beta < \nu)$$

that is on at most a denumerable set.

Suppose that the sequence $A_{\mu(\lambda)}$ is

$$l_1, l_2, l_3, \cdots, l_n, \cdots.$$

Define $f_n(x_\lambda) = l_n$.

Evidently $f_n(x) \to 0$. Moreover, the sequence $f_n(x)$ is uniformly convergent on no non-denumerable set; for if it were uniformly convergent on some non-denumerable set, then, by Auxiliary Theorem I, it would be dominated by a null sequence on this set, which would imply that $B(x)$ is dominated by a null sequence over a non-denumerable set, and this has been shown to be impossible.

1.4. Let us summarize these results in the following manner. In examining the elements of the class Ψ six possibilities present themselves.

(i) The class Ψ is the null set.
(ii) The class Ψ contains only the null set.
(iii) The class Ψ contains only finite subsets of S.
(iv) The class Ψ contains only denumerable subsets of S.
(v) The class Ψ contains at least one non-denumerable subset of S.
(vi) The class Ψ contains the set S.

The results of the preceding section show that:

I. If S is itself denumerable, then (i) and (ii) are impossible, but it is possible for (iii), (iv) and (vi) to occur.

II. If S is itself non-denumerable, then (i), (ii) and (iii) are impossible, but it is possible for (iv), (v) and (vi) to occur.

Hence it appears that in a special case where S is non-denumerable and no U is non-denumerable, the sequence is behaving with the utmost possible stubbornness with respect to the property of uniform convergence.

II. *Approximate uniform convergence.* So far in our investigation of the sets U in the class Ψ we have confined ourselves to their power. We now make a finer distinction and consider their measure.

A sequence $f_n(x)$ is said to be *approximately uniformly convergent* on S if for every positive ϵ there is a U such that $m_e(S - U) < \epsilon$.[9] We would like to find conditions on the set S and the functions $f_n(x)$ which are sufficient to insure the above phenomenon.

[9] We might use a weaker inequality: $m_e(S) - m_e(U) < \epsilon$. As far as I know this does not lead to any results.

A set of sufficient conditions is supplied by the theorem of Egoroff.[10] We have already alluded to this theorem and the work of F. Riesz in deriving simple proofs. Perhaps the simplest of these proofs is that of the Monatshefte.[11] In the light of the first chapter it is in my opinion possible to get a better appreciation of this beautiful proof of Riesz.

As a first trivial remark [12] let us point out that *if all the sets $S - S(k,\nu)$ are closed, then Ψ contains S; i. e., the sequence is uniformly convergent on the whole set S.*

Convergence on the part of the sequence $f_n(x)$ implies that $S - S(k,\nu)$ contains $S - S(k,\nu+1)$, and $\prod_{\nu=1}^{\infty}\{S - S(k,\nu)\} = 0$ [13] for every k. Hence for any particular k it is true, since these sets are all closed, that there exists some ν_k such that $S - S(k,\nu_k) = 0$.

Pick $\nu_1, \nu_2, \nu_3, \cdots$ as the sequence (ν) and consider the set $U(\nu) = \prod_{k=1}^{\infty} S(k,\nu_k)$.

Since $S - U(\nu) = \sum_{k=1}^{\infty}\{S - S(k,\nu_k)\} = 0$ it is clear that the sequence is uniformly convergent on the whole set S.

2. 2. One might make the intuitive remark that *if all the sets $S - S(k,\nu_k)$ are nearly closed then the sequence is uniformly convergent on very nearly the whole set S.*

In this connection let us make the following definition, keeping as close as possible to the familiar notions of classical analysis.

A set S is *approximately closed* [14] if for every $\epsilon > 0$ there exists a set s such that $m_e(s) < \epsilon$ and $S - s$ is closed.

Before proceeding to discuss the above intuitive remark we shall state a few properties of approximately closed sets. All sets mentioned are assumed to be bounded.

2. 3. *The sum of two approximately closed sets is approximately closed.*

2. 4. *A bounded open set is approximately closed.*

2. 5. *The complement of an approximately closed set in an interval is approximately closed.*

---

[10] D. Th. Egoroff, footnote [1], p. 549, *loc. cit.*, p. 244.

[11] F. Riesz, footnote [3] (iii), p. 549, *loc. cit.*, pp. 244-246.

[12] This same trivial remark forms the nucleus of the latest proof by Riesz, footnote [3] (iii), p. 549, *loc. cit.*, p. 244.

[13] $E = 0$ means that E is the null set.

[14] An approximately closed set is clearly measurable, and conversely.

2.6. *The difference of two approximately closed sets is approximately closed.*

2.7. *The product of any number of approximately closed sets is approximately closed.*

2.8. *If the elements in a sequence of approximately closed sets have the character that each contains the next and their product is empty, then their exterior measures approach the limit zero.*

These properties are capable of immediate proof.[15] By way of illustration we shall demonstrate the last.

Suppose that the sets are $S_1, S_2, S_3, \cdots$ and $\prod_1^\infty S_n = 0$. Given any $\epsilon > 0$ pick a series of positive terms $\epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots = \epsilon$. Now $S_n = C_n + s_n$ where C_n is closed and $m_e(s_n) < \epsilon_n$, for every positive integral value of n. Define $C^*_n = C_1 \cdot C_2 \cdot C_3 \cdots C_n$. Then $\prod_1^\infty C^*_n = \prod_1^\infty C_n = 0$, and S_n contains C^*_n.

Since C^*_n contains C^*_{n+1}, and all of these sets are closed, after a certain n_0 all the sets C^*_n are empty. Hence for every $n > n_0$ we have $m_e(C^*_n) = 0$.

If x is not in C^*_n but is in $S_n = S_1 \cdot S_2 \cdot S_3 \cdots S_n$, then x is not in some C_m, $(1 \leq m \leq n)$ but is in all the S_k, $(k = 1, 2, 3, \cdots, n)$. Hence x is in s_m.

Therefore $C^*_n + (s_1 + s_2 + s_3 + \cdots + s_n)$ contains S_n which contains C^*_n and so $S_n = C^*_n + s^*_n$ where $m_e(s^*_n) < \epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots + \epsilon_n$.

For $n > n_0$: $m_e(S_n) \leq m_e(C^*_n) + m_e(s^*_n) < 0 + \epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots + \epsilon_n < \epsilon$.

Hence $m_e(S_n) \to 0$.

2.9. We may now turn to the intuitive remark of 2.2 and show that *if all the sets $S - S(k, \nu)$ are approximately closed then the sequence is approximately uniformly convergent.*[16]

Proof. Given any $\epsilon > 0$ take a convergent series of positive terms $\epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots = \epsilon$.

Since $S - S(k, \nu)$ contains $S - S(k, \nu + 1)$ and $\prod_{\nu=1}^\infty \{S - S(k, \nu)\} = 0$ it follows from 2.8 that $m_e\{S - S(k, \nu)\} \to 0$ as $\nu \to \infty$, for every k.

---

[15] The proofs are materially aided by two criteria: (i) If for every $\epsilon > 0$ there is an s of exterior measure $< \epsilon$ making $S + s$ approximately closed, then S is approximately closed; (ii) If for every ϵ, $S = S_\epsilon + s_\epsilon$ where S_ϵ is approximately closed and $m_e(s_\epsilon) < \epsilon$, then S is approximately closed.

[16] F. Riesz, footnote [3] (iii), p. 549, *loc. cit.*, p. 244.

Thus for every k there exists a ν_k such that

$$m_e\{S - S(k,\nu)\} < \epsilon_k.$$

Define $(\nu) : \nu_1, \nu_2, \nu_3, \cdots$.

Now $m_e\{S - U(\nu)\} = m_e\{S - \prod_{k=1}^{\infty} S(k,\nu_k)\} = m_e\{\sum_{k=1}^{\infty}(S - S(k,\nu_k))\}$
$\leq \sum_{k=1}^{\infty} m_e\{S - S(k,\nu_k)\} \leq \epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots = \epsilon.$

2.10. The problem of insuring approximate uniform convergence on the part of the sequence $f_n(x)$ thus resolves itself into the problem of insuring the sets $S - S(k,\nu)$ to be approximately closed. If we assume that the set S is approximately closed then by 2.6 the problem is further reduced to insuring that the sets $S(k,\nu)$ are approximately closed. This clearly follows if the functions are all continuous, but the condition is much too restrictive. As in the case of sets, we remain as close as possible to the familiar notions of classical analysis and define a class of functions which, while much more general, are close enough to continuous functions to carry us through.

A function $f(x)$ is said to be *approximately continuous* [17] on a set S if for every $\epsilon > 0$ there exists a set s such that $m_e(s) < \epsilon$ and $f(x)$ is continuous on $S - s$ with respect to $S - s$.

We list some properties of approximately continuous functions which lend themselves to immediate proof.

2.11. *The sum, difference, product and quotient (if the denominator is different from zero except on a subset of measure* 0) *of two functions which are approximately continuous on a set S are again approximately continuous on S.*

2.12. *The absolute value of a function which is approximately continuous on S is approximately continuous on S.*

2.13. *If the function $f(x)$ is approximately continuous on an approximately closed set S and if c is any constant, then the subset of S on which the inequality $f(x) \leq c$ holds, is approximately closed.*

2.14. We may now show that *if S is approximately closed and the functions $f_n(x)$ are approximately continuous, then the sets $S - S(k,\nu)$ are approximately closed.*

[17] The identity of functions approximately continuous on an interval and functions measurable on the same interval may be shown by using the theorem of Egoroff, or results of Borel and Hahn, footnote [2] (iii), p. 549, *loc. cit.*, pp. 246-247.

Since S is approximately closed it follows from 2. 6 that it is sufficient to show that the sets $S(k, \nu)$ are approximately closed.

If $s_{k,n,m}$ is the set of points of S where $|f_n(x) - f_m(x)| \leq 1/k$, then $S(k, \nu) = \prod_{n,m=\nu}^{\infty} s_{k,n,m}$. Hence by 2. 7 it remains but to show that the sets $s_{k,n,m}$ are approximately closed. By 2. 11 and 2. 12 $|f_n(x) - f_m(x)|$ is approximately continuous and so by 2. 13 $s_{k,n,m}$ is approximately closed.

2. 15. Combining 2. 9 and 2. 14 we now have the following theorem:

If on an approximately closed set S we have a convergent sequence of approximately continuous functions $f_n(x)$, then the sequence is approximately uniformly convergent.

The theorem is of course valid if the convergence of the hypothesis is convergence almost everywhere. This is the celebrated theorem of Egoroff.

2. 16. An immediate application is the following: *If $f_n(x)$ is a sequence of functions defined and approximately continuous on an approximately closed set S, and if the sequence converges almost everywhere to a function $f(x)$, then $f(x)$ is approximately continuous.*

The proof is immediate by reducing the considerations to a slightly smaller closed set on which all the functions are continuous and the convergence is uniform.

Hence, while in classical analysis it is *not* true that a sequence of *continuous functions* defined and convergent on a *closed set* is *uniformly convergent* and the limit function is *continuous,* the above statement becomes true if the word *approximately* is inserted before the words *continuous, closed* and *uniformly convergent.*

2. 17. It is to be noticed that as a consequence of 2. 5 *a function which is approximately continuous on an approximately closed set may be made approximately continuous on an interval.*

Suppose that $f(x)$ is approximately continuous on an approximately closed set S lying in the interval (a, b). Then the function

$$f^*(x) = \begin{cases} f(x) \text{ on } S \\ 0 \text{ elsewhere in } (a, b) \end{cases}$$

is approximately continuous on (a, b).

It follows that we need consider only functions which are approximately continuous on an interval. We now have the well known theorem:

A necessary and sufficient condition that a function $f(x)$ be approximately continuous on an interval (a, b), is that there exist a sequence of continuous functions defined on (a, b) and converging almost everywhere in this interval to $f(x)$.

The sufficiency follows from 2.16.

To prove the necessity take any convergent series of positive terms $\eta_1 + \eta_2 + \eta_3 + \cdots$. For every n it is true that there is an open set o_n such that $f(x)$ is continuous on $(a, b) - o_n = C_n$ and $m_e(o_n) < \eta_n$. Define a

$$\text{function } c_n(x) = \begin{cases} f(x) \text{ on } C_n \\ \text{linear in the intervals of } o_n \text{ and taking on continuously} \\ \quad \text{the values } f(x) \text{ at the ends of these intervals.} \end{cases}$$

Now $c_n(x)$ is continuous. Moreover, if ξ is in only a finite number of the sets o_n then $c_n(\xi) \to f(\xi)$, since then there is an $n(\xi)$ such that for all $n > n(\xi)$, ξ will be in C_n and so $c_n(\xi) = f(\xi)$.

The set of points in only a finite number of the sets is

$$Z = (\sum_{\nu=1}^{\infty} o_\nu) \cdot (\sum_{\nu=2}^{\infty} o_\nu) \cdot (\sum_{\nu=3}^{\infty} o_\nu) \cdot \cdots.$$

But $\sum_{\nu=n}^{\infty} o_\nu$ contains $\sum_{\nu=n+1}^{\infty} o_\nu$ and $m_e(\sum_{\nu=n}^{\infty} o_\nu) < \eta_n + \eta_{n+1} + \eta_{n+2} + \cdots$. Since the series is convergent $m_e(Z) = 0$ and $c_n(x) \to f(x)$ for almost every x.

This characterization of approximately continuous functions is of immediate importance, since now the theory of Lebesgue integration may be developed along lines similar to the work of Riesz,[18] beginning with the well known concept of the integral of a continuous function. *The salient fact is that the Lebesgue theory of integration may be successfully presented with notions very similar to the familiar concepts of classical analysis* (in fact within epsilon of these).

2.18. Let us come back from the applications of the theorem of Egoroff to the first problem of this chapter. *The problem was to find conditions powerful enough to insure approximate uniform convergence.* One set of conditions has been given by Egoroff, and restricts not only the type of

[18] F. Riesz, footnote [3] (i), p. 549, *loc. cit.*

functions which make up the sequence $f_n(x)$, but also the set S on which they are defined. We find that *it is unnecessary to restrict the set S.* The proof of this fact is based on two auxiliary theorems, the first of which is a lemma of Sierpinski and Zygmund.[19]

AUXILIARY THEOREM I. *If a function $f(x)$ is continuous on a set S then there is an approximately closed set M containing S and a function $f^*(x)$ continuous on M such that $f^*(x) = f(x)$ on S.*

Proof. Consider the closure of S which is $S + S' = S^*$. Let M be the set of points of S^* where the saltus [20] of $f(x)$ is zero. It is clear that M contains S. Moreover, S^* is closed and so the subset S^*_k of its points where the saltus of $f(x)$ is $\geq 1/k$ is closed. But $M = \prod_{k=1}^{\infty} (S^* - S^*_k)$ and so it is approximately closed by 2. 6 and 2. 7. The limit of $f(x)$ at every point of M is unique. Define $f^*(x)$ to be equal to $f(x)$ on S and the limit of $f(x)$ at every point of $M - S$. M is the set and $f^*(x)$ the function required by the theorem.

AUXILIARY THEOREM II. *If a sequence of continuous functions $f_n(x)$ is defined on an approximately closed set M then the subset of M where the sequence converges is approximately closed.*

Proof. Define $\bar{f}_n(x) = \lim_{m\to\infty} \overline{f_n(x), f_{n+1}(x), f_{n+2}(x), \cdots, f_m(x)}$ and $\underline{f}_n(x) = \lim_{m\to\infty} \underline{f_n(x), f_{n+1}(x), f_{n+2}(x), \cdots, f_m(x)}$.[21]

These functions are approximately continuous by 2. 16 for every value of n, and so $\limsup f_n(x) = \lim \bar{f}_n(x)$ and $\liminf f_n(x) = \lim \underline{f}_n(x)$ are also approximately continuous. Hence $(\limsup f_n(x) - \liminf f_n(x))$ is an approximately continuous function by 2. 11, and it follows from 2. 13 that the set of points where the above difference is zero is an approximately closed set. In other words the set of points of M where the limit exists is approximately closed.

We are now in a position to prove that *if a sequence of approximately continuous functions $f_n(x)$ is defined and convergent on any set S, then the sequence is approximately uniformly convergent.*

---

[19] "Sur une fonction discontinue," *Fundamenta Mathematicae*, vol. 4 (1923), p. 317.

[20] The saltus of $f(x)$ at ξ is the least upper bound of the difference $[\limsup f(\dot{x}_n) - \liminf f(\ddot{x}_n)]$ for all possible sequences (chosen from points of S) $\dot{x}_n$ and $\ddot{x}_n$ converging to ξ.

[21] See footnote 6, p. 551.

Proof. Given $\epsilon > 0$ take a sequence of positive terms $\epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots = \epsilon/2$. For every n there is an s_n of exterior measure $< \epsilon_n$ such that $f_n(x)$ is continouus on $S - s_n$. Hence the functions $f_n(x)$, $(n = 1, 2, 3, \cdots)$ are all continuous on the set $S - \Sigma s_n$ and $m_e(\Sigma s_n) < \epsilon/2$.

For every value of n, by Auxiliary Theorem I, there is an approximately closed set M_n containing $S - \Sigma s_n$ and a function $f^*_n(x)$ continuous on M_n such that $f^*_n(x) = f_n(x)$ on $S - \Sigma s_n$. Hence the functions $f^*_n(x)$, $(n = 1, 2, 3, \cdots)$ are all continuous on the approximately closed set $M = \Pi M_n$ and M contains $S - \Sigma s_n$.

By Auxiliary Theorem II the subset M^* of M on which the sequence $f^*_n(x)$ is convergent is approximately closed, and obviously contains $S - \Sigma s_n$.

By the theorem of Egoroff, given $\epsilon/2$ there is a set s of exterior measure $< \epsilon/2$ such that the convergence is uniform on $M^* - s$. Hence the convergence of the sequence $f^*_n(x)$ is uniform, *a fortiori*, on $S - (\Sigma s_n + s)$. Since $f^*_n(x) = f_n(x)$, $(n = 1, 2, 3, \cdots)$ on $S - \Sigma s_n$ and $m_e(\Sigma s_n + s) < \epsilon$ the sequence $f_n(x)$ is approximately uniformly convergent on the set S.

It appears, then, that *when we consider the elements of the class* Ψ *from the point of view of measure we come to the conclusion that as long as the functions of the sequence* $f_n(x)$ *are approximately continuous on* S *there is always a set* U *as close in measure to* S *as we please.*

THE OHIO STATE UNIVERSITY,
COLUMBUS, OHIO.

ON THE MOMENTUM PROBLEM FOR DISTRIBUTION FUNCTIONS IN MORE THAN ONE DIMENSION.

By E. K. HAVILAND.

The Hausdorff momentum problem [1] has recently been solved in the multi-dimensional case by Hildebrandt and Schoenberg.[2] In the present paper there will be treated the corresponding extension of the more general one-dimensional Hamburger [1] momentum problem; i. e., finding a necessary and sufficient condition for the existence of a distribution function [3] ϕ such that

$$c_{nm} = \int\int_S x^n y^m d_{xy}\phi(E), \qquad (n, m = 0, 1, 2, \cdots),$$

where S denotes the entire (x, y)-plane and $\| c_{nm} \|$, $(n, m = 0, 1, 2, \cdots)$, is a given real infinite matrix in which $c_{00} = 1$.

The majority of methods, in particular those based on Jacobi matrices and continued fractions, seem inapplicable in more than one dimension. However, the method of M. Riesz,[4] developed from the ideas of F. Riesz in connection with linear functionals, can be extended to the multi-dimensional case, and the purpose of the present paper is to carry out that extension, the proofs being given, for convenience, in the case of two dimensions.

We consider the operation which makes correspond to any polynomial,

$$P(x, y) = \sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm} x^n y^m,$$

the number

$$P_c = \sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm} c_{nm},$$

where $\| c_{nm} \|$ is a given real infinite matrix. The operation is seen to be distributive. It is said to be non-negative if $P_c \geq 0$ provided $P(x, y)$ is non-negative, i. e., provided $P(x, y) \geq 0$ for all (x, y), and in this case the matrix

[1] For references to literature on the momentum problem in one dimension, cf. M. H. Stone, *op. cit.*, pp. 613-614. References are collected at the end of the present paper.

[2] T. H. Hildebrandt and I. J. Schoenberg, *loc. cit.*

[3] The monotone absolutely additive set function $\phi(E)$ is said to be a distribution function if $0 \leq \phi(E) \leq 1$ and $\phi(S) = 1$, where S denotes the whole (x, y)-plane. Cf. E. K. Haviland, *loc. cit.*, p. 627.

[4] M. Riesz, *loc. cit.*, pp. 4-8.

$\| c_{nm} \|$ will be said to be non-negative. The result of this investigation is then given by the

THEOREM. *For the existence of a distribution function $\phi(E)$ such that*

$$(1) \qquad \iint_S x^n y^m d_{xy}\phi(E) = c_{nm}, \qquad (n, m = 0, 1, 2, \cdots; \; c_{00} = 1),$$

it is necessary and sufficient that the matrix $\| c_{nm} \|$ be non-negative.

Proof. The necessary condition is immediately clear. For if a polynomial $P(x, y) \geqq 0$ for every real (x, y) and if $\phi(E)$ is a distribution function, $\iint_S P(x, y) d_{xy}\phi(E) \geqq 0$ and hence, if (1) is to hold, we must have $P_c \geqq 0$.

To prove the sufficient condition, we let $P_{ij} : (\xi_i, \eta_j)$, $(i, j = 1, 2, \cdots)$ be a denumerable set of points dense in S. In particular, we suppose them to be the intersections of sets of lines parallel to the coördinate axes and everywhere dense in the plane and let the functions $g_{ij}(x, y)$ be defined by $g_{ij}(x, y) = 1$ if $x < \xi_i$ and $y < \eta_j$; while $g_{ij}(x, y) = 0$ otherwise. The operation which makes P_c correspond to the polynomial $P(x, y)$ can be extended to the modul generated by finite linear combinations of $1, x, y, x^2, xy, y^2, \cdots$ $g_{11}(x, y)$, $g_{12}(x, y)$, $g_{21}(x, y), \cdots$ with real constants as coefficients in such a way that the operation remains distributive and non-negative, in the sense that to every non-negative function of this modul there corresponds a non-negative functional value.

This extension is made step by step. We consider first the modul A_1 generated by the various powers $x^m y^n$, $(m, n = 0, 1, 2, \cdots)$, and by $g_{11}(x, y)$. To $g_{11}(x, y)$ we attach, as the value of the functional, a number γ_{11}, attaching at the same time to every finite linear combination of $1, x, y, \cdots$, and $g_{11}(x, y)$ the corresponding combination of c_{00} $(= 1)$, c_{10}, $c_{01}, \cdots$ and γ_{11}. There is thus defined a distributive operation upon the modul A_1.

In order that the operation be non-negative, γ_{11} will have to be so chosen that $\underline{\gamma}_{11} \leqq \gamma_{11} \leqq \bar{\gamma}_{11}$, where $\underline{\gamma}_{11}$ is the upper limit of the values which the operation makes correspond to all polynomials not greater than $g_{11}(x, y)$ for any (x, y) and $\bar{\gamma}_{11}$ is the lower limit of the values which the operation makes correspond to all polynomials not less than $g_{11}(x, y)$ for any (x, y). That such polynomials actually exist may be seen from the fact that $f(x, y) \equiv 0$ belongs to the former class and $f(x, y) \equiv 1$ to the latter. The operation being distributive and non-negative in the modul of polynomials, we shall have [5]

[5] The cases $\underline{\gamma}_{11} = \bar{\gamma}_{11}$ and $\underline{\gamma}_{11} < \bar{\gamma}_{11}$ are associated respectively with the determi-

$\underline{\gamma}_{11} \leq \bar{\gamma}_{11}$. If $\underline{\gamma}_{11} < \bar{\gamma}_{11}$, we choose, for definiteness, $\gamma_{11} = \bar{\gamma}_{11}$. In this way, the operation on the field A_1 is made non-negative as well as distributive, since then $P(x, y) + \mu\, g_{11}(x, y) \geq 0$ implies $P_c + \mu\, \gamma_{11} \geq 0$.

We next form the modul A_2 by adjoining to the generators of A_1 the function $g_{12}(x, y)$. To $g_{12}(x, y)$ we assign as its value the number γ_{12}, the lower limit of the values corresponding to those functions of the modul A_1 which are not less than $g_{12}(x, y)$ for any (x, y). We thus obtain a distributive and non-negative operation defined over the modul A_2. Continuing in this manner, we extend the operation to the moduls $A_3, A_4, \cdots$ obtained by the successive adjunction of the functions $g_{21}(x, y), g_{13}(x, y), \cdots$ to the modul A_2.

We then define a function $F(x, y)$ at the points $P_{ij} : (\xi_i, \eta_j)$ by the equation $F(\xi_i, \eta_j) = \gamma_{ij}$, where γ_{ij} denotes the value of the functional corresponding to the function $g_{ij}(x, y)$. This function $F(x, y)$ possesses, on the points P_{ij}, the monotone property in the sense of Radon.[6] For suppose $\xi_{i_1} < \xi_{i_2}$ and $\eta_{j_1} < \eta_{j_2}$. Then from the definition of the functions $g_{ij}(x, y)$ it follows that for all (x, y)

$$g_{i_1 j_2}(x, y) - g_{i_1 j_1}(x, y) \leq g_{i_2 j_2}(x, y) - g_{i_2 j_1}(x, y).$$

Accordingly, as all these functions are included in some one of the moduls A_k (and, of course, in all succeeding moduls), it follows that

$$0 \leq \gamma_{i_2 j_2} - \gamma_{i_2 j_1} - \gamma_{i_1 j_2} + \gamma_{i_1 j_1},$$

i. e.,

$$0 \leq F(\xi_{i_2}, \eta_{j_2}) - F(\xi_{i_2}, \eta_{j_1}) - F(\xi_{i_1}, \eta_{j_2}) + F(\xi_{i_1}, \eta_{j_1}). \tag{2}$$

Thus $F(x, y)$ possesses on the points P_{ij} the monotone property, q. e. d.

We shall next show that $F(-\infty, y) = F(x, -\infty) = 0$, where $F(-\infty, y) = \lim\limits_{x=-\infty} F(x, y)$, the points (x, y) belonging to the sequence $\{P_{ij}\}$ and the approach to the limit being uniform with respect to y, and a similar interpretation is to be placed on $F(x, -\infty)$. If $\xi_i < 0$, we have $g_{ij}(x, y) \leq \xi_i^{-2} x^2$ for all j and all (x, y). It follows that

$$0 \leq F(\xi_i, \eta_j) = \gamma_{ij} \leq c_{20} \xi_i^{-2}$$

wherefore, as $\xi_i \to -\infty$, $\lim F(\xi_i, \eta_j) = 0$ uniformly for all η_j, q. e. d.

Similarly, if $\eta_j < 0$, we have $g_{ij}(x, y) \leq \eta_j^{-2} y^2$, so it may be shown in a

---

nateness or the non-determinateness of the momentum problem. For the one-dimensional case, cf. M. Riesz, *loc. cit.*, p. 9.

[6] Cf. J. Radon, *loc. cit.* I, p. 1304.

similar manner that as $\eta_j \to -\infty$, $\lim F(\xi_i, \eta_j) = 0$ uniformly for all ξ_i. Again, $g_{ij}(x,y) \leq 1$, wherefore

$$0 \leq F(\xi_i, \eta_j) = \gamma_{ij} \leq c_{00} = 1.$$

With $F(x,y)$ there may be associated an interval function $\psi(I)$ defined for an everywhere dense [7] set of intervals

$$I : (\xi_{i_1} \leq x < \xi_{i_2};\ \eta_{j_1} \leq y < \eta_{j_2})$$

by $$\psi(I) = F(\xi_{i_2}, \eta_{j_2}) - F(\xi_{i_2}, \eta_{j_1}) - F(\xi_{i_1}, \eta_{j_2}) + F(\xi_{i_1}, \eta_{j_1}),$$

and the definition continues to hold when $\xi_{i_1} = \eta_{j_1} = -\infty$, in which case $\psi(I_{i_2 j_2}) = F(\xi_{i_2}, \eta_{j_2})$, where $I_{i_2 j_2} : (-\infty < x < \xi_{i_2};\ -\infty < y < \eta_{j_2})$. From its definition $\psi(I)$ is seen to be additive and we have already seen from equation (2) that it possesses the monotone property. In consequence, as $h \to 0$, $k \to 0$, $\lim F(\xi - h, \eta - k)$ exists, provided $h \geq 0$, $k \geq 0$ and the points $(\xi - h, \eta - k)$, with perhaps the exception of (ξ, η), belong to the everywhere dense set of points for which F is defined.

We now extend the definition of F to all points (x,y) *not* belonging to the given everywhere dense set by setting $F(x,y) = \lim\limits_{h=0,\,k=0} F(x-h, y-k)$, where $(x-h, y-k)$ *is* a point of the everywhere dense set and $h \geq 0$, $k \geq 0$, and we define $\psi(I)$ for *any* interval $I : (x_1 \leq x < x_2;\ y_1 \leq y < y_2)$ by

$$\psi(I) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1). \tag{3}$$

It is seen from the definition that $\psi(I)$ is additive. Moreover, it is monotone, for there exist points (ξ_{i_1}, η_{j_1}), (ξ_{i_1}, η_{j_2}), (ξ_{i_2}, η_{j_1}), (ξ_{i_2}, η_{j_2}) such that for any $x_1 < x_2$ and $y_1 < y_2$

$$0 \leq F(x_m, y_n) - F(\xi_{i_m}, \eta_{j_n}) < \epsilon, \qquad (m, n = 1, 2),$$

and this, together with (2), implies

$$0 \leq F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1).$$

Similarly, it may be shown that $F(x,y)$, as thus defined for all points of the plane, is such that $F(-\infty, y) = F(x - \infty) = 0$, while $F(x,y) \leq 1$ for all (x,y). In fact, it will appear later that $F(+\infty, +\infty) = 1$. Hence $\psi(I)$ is a bounded additive monotone non-decreasing interval function. It follows [8]

[7] Cf. E. K. Haviland, *loc. cit.*, p. 628, Definition 4.

[8] The proof is similar to that given by J. Radon, *loc. cit.* II, p. 1093.

that its discontinuities, if any, fall upon a denumerable set of lines parallel to the coördinate axes. In consequence, there is an everywhere dense set of points (ξ'_i, η'_j), which may be taken to be the intersections of two everywhere dense sets of lines parallel to the coördinate axes, such that

$$\lim_{h=0,\, k=0} F(\xi'_i - h,\ \eta'_j - k) = F(\xi'_i, \eta'_j), \qquad h \geq 0,\ k \geq 0,$$

where the points $(\xi'_i - h,\ \eta'_j - k)$ belong to the same everywhere dense set as does (ξ'_i, η'_j). Then there exists [9] a bounded monotone absolutely additive set function $\phi(E)$ whose corresponding point function, $G(x, y)$, coincides with $F(x, y)$ on the everywhere dense set of points (ξ'_i, η'_j). Moreover, $G(x, y)$ and $F(x, y)$ have the same discontinuity points and are equal at all other points, i. e., $\psi(I)$ and $\phi(E)$ are equal on all their non-singular rectangles. We shall show that $\phi(E)$ is a solution of the momentum problem belonging to the preassigned matrix $\| c_{nm} \|$.

To this end, we consider a monomial [10] $x^n y^m$, where n, m are arbitrary non-negative integers. Let $2r$ be a fixed even number greater than $n + m$, and choose $-T_1 < 0$ and $T_2 > 0$ so that they belong to the set $\xi_1, \xi_2, \cdots$, and $-T'_1 < 0$ and $T'_2 > 0$ so that they belong to the set $\eta_1, \eta_2, \cdots$. Furthermore, T_1, T_2, T'_1, T'_2 shall be so large that

$$| x^n y^m | < \epsilon(x^{2r} + y^{2r})$$

outside the rectangle R : $(-T_1 \leq x < T_2;\ -T'_1 \leq y < T'_2)$ and on its boundary, ϵ being a fixed arbitrarily small positive quantity. We divide R by lines $x_1 = \xi_{i_1} = -T_1$, $x_2 = \xi_{i_2}, \cdots, x_{p+1} = \xi_{i_{p+1}} = T_2$ and $y_1 = \eta_{j_1} = -T'_1$, $y_2 = \eta_{j_2}, \cdots, y_{q+1} = \eta_{j_{q+1}} = T'_2$ into a set of rectangles

$$R_{kl} : (x_k \leq x < x_{k+1};\ y_l \leq y < y_{l+1})$$

in each of which the oscillation of $x^n y^m$ is less than ϵ', where ϵ' is another fixed arbitrarily small positive quantity. Let (X_k, Y_l) be a point in the interior of the rectangle R_{kl}. We then form the step function $v(x, y)$ which vanishes outside R and which takes the value $X_k{}^n Y_l{}^m$ in R_{kl}. Then for every (x, y)

$$v(x, y) - \epsilon' - \epsilon(x^{2r} + y^{2r}) < x^n y^m < v(x, y) + \epsilon' + \epsilon(x^{2r} + y^{2r}).$$

Since the function $v(x, y)$ belongs to one of the moduls $A_1, A_2, \cdots$ (and

[9] Cf. E. K. Haviland, *loc. cit.*, p. 651, and the references there given.

[10] The proof holds also for an arbitrary polynomial.

hence to any subsequent modul), the functional operation is defined for it, and if to $v(x, y)$ corresponds the functional value v_c,

$$v_c - \epsilon' - \epsilon(c_{2r,0} + c_{0,2r}) \leq c_{nm} \leq v_c + \epsilon' + \epsilon(c_{2r,0} + c_{0,2r}). \tag{4}$$

Furthermore,

$$v(x,y) = \sum_{k=1}^{p} \sum_{l=1}^{q} X_k{}^n Y_l{}^m [g_{i_{k+1}j_{l+1}}(x,y) - g_{i_{k+1}j_l}(x,y) - g_{i_k j_{l+1}}(x,y) + g_{i_k j_l}(x,y)].$$

Hence, as $\gamma_{i_k j_l} = F(\xi_{i_k}, \eta_{j_l}) = F(x_k, y_l)$, we have by (3)

$$v_c = \sum_{k=1}^{p} \sum_{l=1}^{q} X_k{}^n Y_l{}^m \psi(R_{kl}).$$

The inequality (4) can then be written

$$\begin{aligned} \sum_{k=1}^{p} \sum_{l=1}^{q} X_k{}^n Y_l{}^m \psi(R_{kl}) - \epsilon' - \epsilon(c_{2r,0} + c_{0,2r}) &\leq c_{nm} \\ &\leq \sum_{k=1}^{p} \sum_{l=1}^{q} X_k{}^n Y_l{}^m \psi(R_{kl}) + \epsilon' + \epsilon(c_{2r,0} + c_{0,2r}). \end{aligned} \tag{5}$$

Now $x^n y^m$ is continuous in x and y together in every rectangle and $\psi(I)$ is a bounded monotone additive interval function. Hence,[11] as the diameter of the R_{kl} approaches zero,

$$\lim \sum_{k=1}^{p} \sum_{l=1}^{q} X_k{}^n Y_l{}^m \psi(R_{kl}) = \int\int_R x^n y^m d_{xy}\psi(I).$$

At the same time, $\epsilon' \to 0$, so that from (5) we obtain

$$\begin{aligned} \int\int_R x^n y^m d_{xy}\psi(I) - \epsilon(c_{2r,0} + c_{0,2r}) \\ \leq c_{nm} \leq \int\int_R x^n y^m d_{xy}\psi(I) + \epsilon(c_{2r,0} + c_{0,2r}). \end{aligned}$$

Let $T_1, T'_1, T_2, T'_2 \to +\infty$ and $\epsilon \to 0$. Then

$$\int\int_S x^n y^m d_{xy}\psi(I) = c_{nm}.$$

As, however, for any arbitrarily large non-singular rectangle R_1 of ϕ and ψ,

$$\int\int_{R_1} x^n y^m d_{xy}\psi(I) = \int\int_{R_1} x^n y^m d_{xy}\phi(E),$$

it follows that[12]

---

[11] Cf., e. g., S. Bochner, *loc. cit.*, p. 391.

[12] This follows directly if m and n are both even. Otherwise one has first to use

$$\iint_S x^n y^m d_{xy}\phi(E) = c_{nm}, \qquad \text{q. e. d.}$$

The distribution function ϕ whose existence is thus established is not necessarily uniquely determined. Sufficient conditions that ϕ be uniquely determined by its momenta c_{nm} have been found by V. Romanovsky [13] and by the present author.[14] In particular, ϕ is uniquely determined if the c_{nm} are such that

$$\left(\sum_{\nu=0}^{2n} \binom{2n}{\nu} \mid c_{2n-\nu,\nu} \mid \right)^{1/(2n)} = o(n),$$

and the present author has shown [14] that this condition is almost necessary in that ϕ may not be uniquely determined if $o(n)$ is replaced by $o(n^{1+\epsilon})$.

REFERENCES.

S. Bochner, "Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse," *Mathematische Annalen*, vol. 108 (1933), pp. 378-410.

E. K. Haviland, "On the theory of absolutely additive distribution functions," *American Journal of Mathematics*, vol. 56 (1934), pp. 625-658.

T. H. Hildebrandt and I. J. Schoenberg, "On linear functional operators and the moment problem for a finite interval in one or several dimensions," *Annals of Mathematics*, ser. 2, vol. 34 (1933), pp. 317-328.

J. Radon, I: "Theorie und Anwendungen der absolut additiven Mengenfunktionen," *Sitzungsberichte der mathematischen-naturwissenschaftlichen Klasse der Kaiserl. Akademie zu Wien*, vol. 122 (1913), pp. 1295-1438; II: "Uber lineare Funktionaltransformationen und Funktionalgleichungen," *ibid.*, vol. 128 (1919), pp. 1083-1121.

M. Riesz, "Sur le problème des moments," *Arkiv för Matematik, Astronomi och Fysik*, vol. 17 (1923), No. 16.

V. Romanovsky, "Sur un théorème limite du calcul des probabilités," *Recueil Mathématique de la Société Mathématique de Moscou*, vol. 36 (1929), pp. 36-64.

M. H. Stone, "Linear transformations in Hilbert space," *American Mathematical Society Colloquium Publications*, vol. 15.

THE JOHNS HOPKINS UNIVERSITY.

---

the inequality of Schwarz. For the existence of the integral with respect to ϕ, cf. J. Radon, *loc. cit.* I, pp. 1322-1324. Notice that if $n = m = 0$, we obtain $\phi(S) = c_{00} = 1$.

[13] V. Romanovsky, *loc. cit.*, p. 47, § 3.

[14] E. K. Haviland, *loc. cit.*, p. 634. At the time of publishing that paper, the author was not aware that a proof of his Theorem I, under restricted conditions, and of the sufficient condition in his Theorem II had previously been given by Romanovsky. Romanovsky's statement of this sufficient condition differs somewhat from the present author's but the two proofs are effectively the same.

A NOTE ON A PROPERTY OF FOURIER-STIELTJES TRANSFORMS IN MORE THAN ONE DIMENSION.

By E. K. Haviland.

If $\rho(\xi)$ is monotone in $[-\infty, +\infty]$ and $\rho(-\infty) = 0$, $\rho(+\infty) = 1$; if $\Lambda(t;\rho) = \int_{-\infty}^{+\infty} \exp(it\xi)\, d\rho(\xi)$, where $-\infty < t < +\infty$, denotes the Fourier-Stieltjes transform of ρ; and if, finally, $\mathfrak{M}(f(\cdot)) = \lim_{T=+\infty} (2T)^{-1} \int_{-T}^{T} f(t)\, dt$, it is known [1] that $\mathfrak{M}(|\Lambda(\cdot\,;\rho)|^2) = \Sigma\, |\Delta_k|^2$ where $\Delta_k = \rho(\xi_k + 0) - \rho(\xi_k - 0)$ and the summation is taken over all the (at most denumerable) discontinuities of ρ. In particular, if ρ is continuous,

$$\mathfrak{M}(|\Lambda(\cdot\,;\rho)|^2) = 0. \tag{1}$$

In the case of more than one dimension, the discontinuity points need no longer be denumerable, so the question arises as to what then corresponds to the foregoing result. It turns out that in the multi-dimensional case a similar result holds, the point spectrum, as defined in an earlier paper,[2] playing the rôle of the discontinuities in the one-dimensional case, while the "mild" discontinuity points, i. e., those not occurring in the point spectrum, play no rôle at all. More precisely, it will be shown that in more than one dimension,[3] if $\phi(E)$ be a distribution function,[4]

---

[1] This result was stated, without proof, by Paul Lévy, *Calcul de probabilités* (Paris, 1925), p. 171. For a proof, cf. I. Schoenberg, "Über total monotone Folgen mit stetiger Belungungsfunktion," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 761-767, where reference is made to a paper of N. Wiener. Since then, (1) has often been rediscovered in connection with the unitary dynamics of Carleman and Koopman and with the statistical considerations of Khintchine. Cf. also A. Wintner and E. K. Haviland, "On the Fourier-Stieltjes transform," *American Journal of Mathematics*, vol. 56 (1934), pp. 4-5.

[2] Cf. E. K. Haviland, "On the theory of absolutely additive distribution functions," *American Journal of Mathematics*, vol. 56 (1934), p. 654.

[3] For convenience, we give the proof in the two-dimensional case.

[4] The monotone absolutely additive set function $\phi(E)$ is said to be a distribution function if $0 \leq \phi(E) \leq 1$ and $\phi(S) = 1$, where S denotes the whole plane. Cf. E. K. Haviland, *ibid.*, p. 627. For the definition of integrals with respect to such functions, cf. J. Radon, "Theorie und Anwendungen der absolut additiven Mengenfunktionen," *Sitzungsberichte der mathematischen-naturwissenschaftlichen Klasse der Kaiserl. Akademie zu Wien*, vol. 122 (1913), pp. 1322-1324.

(2) $$\mathfrak{M}(|\Lambda(s,t;\phi)|^2) = \sum_{k=1}^{\infty} [\phi(P_k)]^2,$$

where

(3) $$\Lambda(s,t;\phi) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\{i(sx+ty)\} d_{xy}\phi(E)$$

and

$$\mathfrak{M}(f(s,t)) = \lim_{T=\infty,\, U=\infty} (2T)^{-1}(2U)^{-1} \int_{-T}^{T}\int_{-U}^{U} f(s,t)\, ds\, dt,$$

and the summation on the right of (2) is taken over all points P_k of the point spectrum of ϕ.

While the result (2) is analogous to that in the one-dimensional case, it is not obvious from the latter, since in two or more dimensions the singularities of a monotone function are essentially more complicated than those of such a function in a single dimension, where *all* discontinuity points belong to the point spectrum.

The proof of (2) is as follows: If $\phi(E)$ be a distribution function, we define a set function $\bar{\phi}(E)$ by setting

(4) $$\bar{\phi}(E) = \phi(-E),$$

where $-E$ is the set symmetric to E with respect to the origin. Then $\bar{\phi}(E)$ is a distribution function, and,[5] by virtue of the Convolution Theorem for Fourier-Stieltjes transforms,[6] $\Lambda(s,t;\phi * \bar{\phi}) = \Lambda(s,t;\phi) \cdot \Lambda(s,t;\bar{\phi})$; or, since $\Lambda(s,t;\phi)$ and $\Lambda(s,t;\bar{\phi})$ are conjugated complex quantities in virtue of (3) and (4),

$$|\Lambda(s,t;\phi)|^2 = \Lambda(s,t;\phi * \bar{\phi}).$$

Consequently

(5) $$\mathfrak{M}(|\Lambda(s,t;\phi)|^2) = \mathfrak{M}(\Lambda(s,t;\phi * \bar{\phi}))$$

and we need examine only the latter.

It is now to be shown that if $\psi(E)$ be a distribution function whose point spectrum is vacuous, then $\mathfrak{M}(\Lambda(s,t;\psi)) = 0$. Since the contribution of the integration domain $S - R$ to $\Lambda(s,t;\psi)$, where S represents the entire (x,y)-plane and R an arbitrary rectangle in that plane having its sides parallel to the coördinate axes, is in absolute value less than ϵ for all (s,t) provided R is sufficiently large, it is sufficient to prove that for any fixed R and for sufficiently large values of T, U

---

[5] $\psi_1 * \psi_2$ denotes the symbolical product (Faltung or convolution) of ψ_1 and ψ_2. Cf. E. K. Haviland, *loc. cit.*, p. 651, Theorem IV.

[6] Cf. E. K. Haviland, *ibid.*, Theorem V.

$$\left| (4TU)^{-1} \int_{-T}^{T} \int_{-U}^{U} \left\{ \int\int_{R} \exp[i(sx+ty)] d_{xy}\psi(E) \right\} ds\, dt \right| < \epsilon.$$

We begin the proof of this statement by observing that the expression beneath the absolute value signs may be written as

$$J = (4TU)^{-1} \int_{-T}^{T} \int_{-U}^{U} \left\{ \int_{-M}^{M} \int_{-N}^{N} \exp[i(sx+ty)] d_{xy}\psi(E) \right\} ds\, dt,$$

and it is permissible to invert the order of integration,[7] obtaining

$$\begin{aligned} J &= \int_{-M}^{M} \int_{-N}^{N} \left\{ (2T)^{-1} \int_{-T}^{T} e^{isx} ds \cdot (2U)^{-1} \int_{-U}^{U} e^{ity} dt \right\} d_{xy}\psi(E) \\ &= \int_{-M}^{M} \int_{-N}^{N} \frac{\sin Tx}{Tx} \frac{\sin Uy}{Uy} d_{xy}\psi(E) \\ &= \int\int_{\mathrm{I}} + \int\int_{\mathrm{II}} + \int\int_{\mathrm{III}} + \int\int_{\mathrm{IV}} + \int\int_{\mathrm{V}}, \end{aligned}$$

where

$\mathrm{I}: (-M \leq x < -\delta; -N \leq y \leq N)$, $\mathrm{II}: (-\delta \leq x \leq \delta; \delta < y \leq N)$,
$\mathrm{III}: (\delta < x \leq M; -N \leq y \leq N)$, $\mathrm{IV}: (-\delta \leq x \leq \delta; -N \leq y < -\delta)$,
$\mathrm{V}: (-\delta \leq x \leq \delta; -\delta \leq y \leq \delta)$.

Now $\left| \int\int_{\mathrm{V}} \right| \leq \int\int_{\mathrm{V}} d_{xy}\psi(E)$ and if $(0,0)$ is not a point of the point spectrum of ψ, this last expression may be made less than $\epsilon/5$ in absolute value by taking δ sufficiently small. δ being thus fixed, it is easily seen that

$$\left| \int\int_{\mathrm{I}} \right| \leq (T\delta)^{-1} \int\int_{\mathrm{I}} d_{xy}\psi(E) < \epsilon/5$$

and

$$\left| \int\int_{\mathrm{III}} \right| \leq (T\delta)^{-1} \int\int_{\mathrm{III}} d_{xy}\psi(E) < \epsilon/5$$

provided T is sufficiently large. Similarly,

$$\left| \int\int_{\mathrm{II}} \right| \leq (U\delta)^{-1} \int\int_{\mathrm{II}} d_{xy}\psi(E) < \epsilon/5$$

and

$$\left| \int\int_{\mathrm{IV}} \right| \leq (U\delta)^{-1} \int\int_{\mathrm{IV}} d_{xy}\psi(E) < \epsilon/5$$

if U is sufficiently large. Consequently, if the point spectrum of ψ is vacuous, $\mathfrak{M}(\Lambda(s,t;\psi)) = 0$.

[7] Cf. E. K. Haviland, *ibid.*, p. 640.

Furthermore,[8] every absolutely additive set function ϕ of bounded total variation is the sum of two functions, say ϕ_1 and ϕ_2, of which the former has a vacuous point spectrum, while the latter is purely discontinuous (i. e., its spectrum coincides with its point spectrum). Then from the definition of $\bar{\phi}$ it follows that $\bar{\phi} = \bar{\phi}_1 + \bar{\phi}_2$ and

$$\begin{aligned}\mathfrak{M}(\Lambda(s,t;\phi * \bar{\phi})) &= \mathfrak{M}(\Lambda(s,t;(\phi_1 + \phi_2) * (\bar{\phi}_1 + \bar{\phi}_2)) \\ &= \mathfrak{M}(\Lambda(s,t;\phi_1 * \bar{\phi}_1)) + \mathfrak{M}(\Lambda(s,t;\phi_1 * \bar{\phi}_2)) \\ &\quad + \mathfrak{M}(\Lambda(s,t;\bar{\phi}_1 * \phi_2)) + \mathfrak{M}(\Lambda(s,t;\phi_2 * \bar{\phi}_2)).\end{aligned}$$

But the point spectra of $\phi_1 * \bar{\phi}_1$, $\phi_1 * \bar{\phi}_2$ and $\bar{\phi}_1 * \phi_2$ are vacuous by the addition rule [9] for point spectra, so that the first three terms in the last member of the preceding equation vanish, and

$$\begin{aligned}(6)\qquad \mathfrak{M}(\Lambda(s,t;\phi * \bar{\phi})) &= \mathfrak{M}(\Lambda(s,t;\phi_2 * \bar{\phi}_2)) \\ = \mathfrak{M}(\Lambda(s,t;\phi_2)\cdot\Lambda(s,t;\bar{\phi}_2)) &= \mathfrak{M}(\Lambda(s,t;\phi_2)\cdot\Lambda(-s,-t;\phi_2)).\end{aligned}$$

Let points of the point spectrum of ϕ, i. e., of ϕ_2, be $P_k : (x_k, y_k)$. As they are at most denumerable, it follows that

$$\Lambda(s,t;\phi_2) = \sum_{k=1}^{\infty} \exp[i(sx_k + ty_k)]\phi(P_k)$$

and

$$\Lambda(-s,-t;\phi_2) = \sum_{k=1}^{\infty} \exp[-i(sx_k + ty_k)]\phi(P_k).$$

Substituting in (6) and taking account of (5), we find

$$\mathfrak{M}(|\Lambda(s,t;\phi)|^2) = \mathfrak{M}(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \exp[i\{s(x_j - x_k) + t(y_j - y_k)\}]\phi(P_j)\phi(P_k)).$$

On taking the mean value, all terms for which $j \neq k$ disappear, so that the right-hand side of the preceding equation becomes $\sum_{k=1}^{\infty} [\phi(P_k)]^2$ which proves (2).

THE JOHNS HOPKINS UNIVERSITY.

[8] Cf. H. Hahn, *Theorie der reelen Funktionen* (Berlin, 1921), p. 414, Theorem XV.

[9] Cf. E. K. Haviland, *loc. cit.*, p. 654, Theorem VI.

THE THEORY OF THE SECOND VARIATION FOR THE NON-PARAMETRIC PROBLEM OF BOLZA.[1]

By William T. Reid.

1. **Introduction.** The non-parametric problem of Bolza in the calculus of variations is that of finding in a class of arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leq x \leq x_2), \tag{1.1}$$

satisfying the differential equations and end conditions

$$\phi_\alpha[x, y, y'] = 0 \qquad (\alpha = 1, \cdots, m < n), \tag{1.2}$$

$$\psi_\gamma[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\gamma = 1, \cdots, p \leqq 2n + 2), \tag{1.3}$$

one which minimizes an expression of the form

$$I = G[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f(x, y, y')\,dx. \tag{1.4}$$

Sufficient conditions for the problem of Bolza have been given by Morse [III][2] and Bliss [IV] for extremal arcs that are not only normal relative to the end conditions but also normal on sub-intervals. Recently, sufficient conditions have been obtained under weaker assumptions by Hestenes [X], who has replaced the usual condition of Mayer by a new condition in terms of a certain quadratic form involving the solutions of the accessory equations. Hestenes has not only been able to discard the hypothesis of normality on sub-intervals, but has also obtained sufficient conditions for an extremal arc with multipliers of the form $\lambda_0 = 1$, $\lambda_\alpha(x)$ which is not necessarily normal relative to the end conditions.

The principal result of the present paper is the following theorem:

Theorem A. *If E_{12} : $y_i = y_i(x)$, $\lambda_0 =$ constant, $\lambda_\alpha(x)$, $x_1 \leq x \leq x_2$, is an extremal arc which satisfies the strengthened Clebsch condition, and on which there is no point conjugate to the point 1, then there exists a family*

[1] Presented to the American Mathematical Society, September 5, 1934.

[2] Roman numerals in brackets refer to the bibliography at the end of this paper. Only papers to which direct reference is made in the present paper are listed. For a more extensive bibliography the reader is referred to that given by Hestenes at the end of [X].

of n mutually conjugate accessory extremals $\eta_{ij}(x)$, $\zeta_{ij}(x)$ $(j=1,\cdots,n)$ *such that* $|\eta_{ij}(x)|\neq 0$ *on* x_1x_2.

This theorem is fundamental in the construction of a field of extremals imbedding a given extremal, and has been proved by several authors under additional assumptions of normality.[3] Theorem A has been proved by Morse by an extension of the methods which he used in [II].[4] *The chief significance of the independent proof here given is that it is a direct generalization of the method used by Bliss when the extremal arc satisfies additional normality conditions* [I, pp. 729, 736], *and hence is more intimately related to the methods usually used in the simpler problems of the calculus of variations than the methods of Morse and Hestenes.*

Certain general properties of accessory extremals are discussed in § 2 of this paper, and Theorem A is established in § 3. In § 4 there are proved, by the use of the results of § 3, further results concerning the existence of families of accessory extremals satisfying the condition of Theorem A. *In particular, Theorem* 4. 2 *gives a rather elegant method for determining such a family of accessory extremals.* It is to be noted, however, that the proof of the interesting result of Theorem 4. 3 is independent of the results of § 3. Finally, in § 5 there is discussed briefly the relation of Theorem A to sufficiency theorems for the problems of Bolza and Mayer.

Throughout the paper, the coefficients of (1. 2), (1. 3), and (1. 4) are supposed to satisfy the hypotheses usually made {see [III], [IV] and [X]}.

2. Accessory extremals. For an extremal arc $E_{12}: y_i = y_i(x)$, $\lambda_0 =$ constant, $\lambda_\alpha(x)$, $x_1 \leq x \leq x_2$, let

$$
\begin{aligned}
&F = \lambda_0 f(x,y,y') + \lambda_\alpha(x)\phi_\alpha(x,y,y'),\\
(2.1)\quad &2\omega[x,\eta,\eta'] = F_{y'_iy'_j}\eta'_i\eta'_j + 2F_{y'_iy_j}\eta'_i\eta_j + F_{y_iy_j}\eta_i\eta_j,\\
&\Phi_\alpha[x,\eta,\eta'] = \phi_{\alpha y'_i}\eta'_i + \phi_{\alpha y_i}\eta_i \qquad (\alpha=1,\cdots,m).
\end{aligned}
$$

The coefficients of ω and ϕ_α are supposed to have as arguments the functions $y_i(x)$, λ_0, $\lambda_\alpha(x)$ belonging to E_{12}. It will also be supposed that E_{12} satisfies the strengthened Clebsch condition [IV, p. 264]. As usual, this condition will be denoted as III′. If we set

[3] See [I], pp. 729, 736; [II]; [VI], p. 320; and [X], pp. 804, 807.

[4] I was not aware that Morse had proved this result until the date upon which I presented my proof to the American Mathematical Society. Morse's paper has since appeared in the *Transactions of the American Mathematical Society*, vol. 37 (1935), pp. 147-160. Hestenes has informed me that subsequent to my proof of Theorem A, he also proved this result by the use of the formulation of the Mayer condition which he has used in [X].

(2.2) $$\Omega[x, \eta, \eta', \mu] = \omega[x, \eta, \eta'] + \mu_a \Phi_a[x, \eta, \eta'],$$

then the system of accessory differential equations is

(2.3) $$(d/dx)\Omega_{\eta'_i} - \Omega_{\eta_i} = 0, \quad \Phi_a = 0 \quad (i = 1, \cdots, n;\ \alpha = 1, \cdots, m).$$

By the introduction of the canonical variables $\zeta_i = \Omega_{\eta'_i}[x, \eta, \eta', \mu]$, this system is seen to be equivalent to a system of $2n$ linear differential equations of the first order of the form

(2.3′) $$\eta'_i = A_{ij}(x)\eta_j + B_{ij}(x)\zeta_j, \quad \zeta'_i = C_{ij}(x)\eta_j - A_{ji}(x)\zeta_j.$$

The coefficients in (2.3′) are continuous on x_1x_2, $\| B_{ij} \|$ and $\| C_{ij} \|$ are symmetric matrices, and $\| B_{ij} \|$ is of rank $n - m$ on x_1x_2.[5] We shall say that a set of functions $\eta_i(x)$, $\zeta_i(x)$ which are of class C', and which satisfy (2.3′) on x_1x_2 is an *accessory extremal.*[6]

If the extremal E_{12} is a minimizing arc which is normal on every sub-interval x_3x_2 of x_1x_2 it has been shown by Bliss that there can be no point conjugate to 1 on E_{12} between 1 and 2.[7] We shall use IV_0 to denote this necessary condition, and IV'_0 to denote the condition that there is no value x_3 such that $x_1 < x_3 \leq x_2$ and defining a point 3 conjugate to 1.

We shall say that the order of anormality[8] of E_{12} on a sub-interval t_1t_2 of x_1x_2 is equal to r if on this sub-interval there are exactly r linearly independent accessory extremals $\eta_i = u_{ik}(x)$, $\zeta_i = v_{ik}(x)$ $(k = 1, \cdots, r)$ with $u_{ik}(x) \equiv 0$ on t_1t_2.

The following properties of accessory extremals will be given without proof.

PROPERTY 1°. *The order of anormality of E_{12} on a given sub-interval is at most m.*

---

[5] See, for example, [VIII], §§ 3 and 4.

[6] The terminology *accessory differential system* for the system (2.3) is due to von Escherich. The problem of minimizing the second variation in a class of arcs satisfying the equations of variation has been called the *accessory minimum problem* [III, and X], and the associated boundary value problem has been termed the *accessory boundary value problem* [III, VIII and X]. On the other hand, a set of functions $\eta_i(x)$ belonging to a solution η_i, μ_a of (2.3), or to a solution η_i, ζ_i of (2.3′), has been called a *secondary extremal.* [III, and X]. It seems more consistent to either speak of an *accessory extremal,* or else to use the terms *secondary differential system, secondary minimum problem, secondary boundary value problem,* and *secondary extremal.* Due to the priority of the term *accessory* for the differential system, the present author has adopted the phrase *accessory extremal* in the sense defined above.

[7] See [I], p. 725. The reader is referred to [I] for the definition of conjugate point.

[8] This terminology has been used by Hestenes, [X], p. 799.

PROPERTY 2°. *If the order of anormality of E_{12} on a sub-interval t_1t_2 is r, and $\eta_i \equiv 0$, $\zeta_i = v_{ik}(x)$ $(k = 1, \cdots, r)$ are linearly independent accessory extremals on this sub-interval, then for arbitrary admissible variations $\eta_i(x)$ and arbitrary points x', x'' of t_1t_2, we have*

$$v_{ik}(x)\eta_i(x) \Big|_{x=x'}^{x=x''} = 0 \qquad (k = 1, \cdots, r).$$

We shall denote by $r(x)$ the order of anormality of E_{12} on the sub-interval x_1x $(x_1 < x \leq x_2)$ of x_1x_2. The function $r(x)$ is seen to be monotone non-increasing on $x_1 < x \leq x_2$. In view of Property 1°, we have

PROPERTY 3°. *There exists a constant d such that $0 < d \leq x_2 - x_1$, and $r(x)$ is constant on $x_1 < x \leq x_1 + d$.*

The following property is a consequence of the continuity of the solutions η_i, ζ_i of (2. 3′):

PROPERTY 4°. *If $x_1 < x_3 \leq x_2$, there exists a δ such that $0 < \delta < x_3 - x_1$ and $r(x) = r(x_3)$ on $x_3 - \delta \leq x \leq x_3$.*

In view of the above properties, it is seen that $r(x)$ has at most m points of discontinuity on $x_1 < x \leq x_2$. We shall denote these points by $t_1, \cdots, t_g$, where $x_1 < t_g < t_{g-1} < \cdots < t_1 < x_2$. For convenience, let $t_{g+1} = x_1$, $t_0 = x_2$, and $r_q = r(t_q)$ $(q = 0, 1, \cdots, g)$.

It is readily seen that one may choose a family of accessory extremals $u_{ij}(x)$, $v_{ij}(x)$ $(j = 1, \cdots, n)$ such that

$$(2.4)\quad \begin{aligned} &u_{ij}(x_1) = 0,\ u_{i\theta}(x) \equiv 0 \text{ on } x_1 \leq x \leq t_q \text{ for} \\ &\qquad\qquad\qquad\qquad \theta = 1, \cdots, r_q;\ q = 0, 1, \cdots, g, \\ &v_{ki}(x_1)v_{kj}(x_1) = \delta_{ij}\ (\delta_{ii} = 1,\ \delta_{ij} = 0 \text{ if } i \neq j)\ (i, j = 1, \cdots, n). \end{aligned}$$

Now define another set of n accessory extremals $u_{i\,n+j}(x)$, $v_{i\,n+j}(x)$ by the initial conditions

$$(2.5)\qquad u_{i\,n+j}(x_1) = v_{ij}(x_1),\quad v_{i\,n+j}(x_1) = 0 \qquad (i, j = 1, \cdots, n).$$

Finally, define $u_{ij}(x \mid q)$, $v_{ij}(x \mid q)$ $(q = 0, 1, \cdots, g)$ as follows:

$$(2.6)\quad \begin{aligned} &u_{ij}(x \mid q) = u_{i\,n+j}(x),\ v_{ij}(x \mid q) = v_{i\,n+j}(x) \text{ for } j = 1, \cdots, r_q, \\ &u_{ij}(x \mid q) = u_{ij}(x),\ v_{ij}(x \mid q) = v_{ij}(x) \text{ for } j = r_q + 1, \cdots, n. \end{aligned}$$

The following property follows readily from the definition of a conjugate point:

PROPERTY 5^0. *A value x_3 on the sub-interval $t_{q+1} < x \leq t_q$ $(q = 0, 1, \cdots, g)$ defines a point on E_{12} conjugate to the point 1 if and only if one of the following conditions is satisfied:*

(α) *the matrix* $\left\| \begin{matrix} u_{is}(x_1) \\ u_{is}(x_3) \end{matrix} \right\|$ $(s = 1, \cdots, 2n)$ *has rank less than* $2n - r_q$.

(β) *the matrix* $\| u_{im}(x_3) \| \equiv \| u_{im}(x_3 \mid q) \|$ $(m = r_q + 1, \cdots, n)$ *has rank less than* $n - r_q$.

(γ) *the determinant* $| u_{ij}(x_3 \mid q) |$ *is equal to zero.*

It is verified readily that the accessory extremals of each of the n parameter families defined by (2.4), (2.5) or (2.6) are mutually conjugate [see I, p. 738] in pairs.

3. Proof of Theorem A. For clarity, the essential steps in the proof of this theorem will be stated in the form of lemmas.

LEMMA 3.1. *In terms of the accessory extremals defined by* (2.4) *and* (2.5), *define n mutually conjugate accessory extremals $U_{ij}(x;\rho)$ $V_{ij}(x;\rho)$ as follows:*

$$
\begin{aligned}
&U_{ij}(x;\rho) = u_{i\,n+j}(x) + \rho u_{ij}(x); \; V_{ij}(x;\rho) = v_{i\,n+j}(x) + \rho v_{ij}(x), \\
(3.1) \qquad &\qquad\qquad\qquad\qquad\qquad\qquad \text{for } i = 1, \cdots, n; \; j = 1, \cdots, r_g; \\
&U_{ij}(x;\rho) = u_{ij}(x), \; V_{ij}(x;\rho) = v_{ij}(x) \text{ for } j = r_g + 1, \cdots, n).
\end{aligned}
$$

Then if E_{12} is an extremal satisfying the conditions of Theorem A, the determinant $| U_{ij}(x;\rho) |$ is different from zero on $x_1 < x \leq x_2$ for ρ positive and sufficiently large in value.

This lemma will be proved by induction. In view of condition (γ) of Property 5^0 of § 2, it is seen that $| u_{ij}(x \mid q) | \neq 0$ on $t_{q+1} < x \leq t_q$ $(q = 0, 1, \cdots, g)$. Now for $x_1 < x \leq t_g$ we have $U_{ij}(x;\rho) \equiv u_{ij}(x \mid g)$, and hence for arbitrary values ρ we have $| U_{ij}(x;\rho) | \neq 0$ on $x_1 < x \leq t_g$.

It will now be proved that if for a value $q = \sigma$ there exists a positive value $\rho = \rho_1$ such that $| U_{ij}(x;\rho_1) | \neq 0$ on $x_1 < x \leq t_\sigma$, then there exists a value $\rho_2 > \rho_1$ such that $| U_{ij}(x;\rho_2) | \neq 0$ on $x_1 < x \leq t_{\sigma-1}$. By hypothesis, $| U_{ij}(x;\rho_1) | \neq 0$ on $x_1 < x \leq t_\sigma$. Hence there exists an ϵ such that $t_\sigma < t_\sigma + \epsilon < t_{\sigma-1}$ and $| U_{ij}(x;\rho_1) | \neq 0$ on $x_1 < x \leq t_\sigma + \epsilon$. It will first be shown that if $\rho > \rho_1$, then $| U_{ij}(x;\rho) | \neq 0$ also on $x_1 < x \leq t_\sigma + \epsilon$. For if there were a value x_3 on this interval such that $| U_{ij}(x_3;\rho) | = 0$, there would exist constants c_j not all zero and such that $U_{ij}(x_3;\rho) c_j = 0$. For these constants, let

(3.2) $\qquad \eta_i(x) = U_{ij}(x;\rho)c_j, \quad \zeta_i(x) = V_{ij}(x;\rho)c_j \qquad (x_1 \leq x \leq x_3).$

Since on $x_1 t_g$ we have $U_{ij}(x;\rho) \equiv U_{ij}(x;\rho_1)$, and $| U_{ij}(x;\rho_1) | \neq 0$ on $x_1 < x \leq t_\sigma + \epsilon$, it would follow that the functions $\eta_i(x)$ defined by (3.2) are of the form $\eta_i(x) = U_{ij}(x;\rho_1)a_j(x)$ on $x_1 \leq x \leq x_3$, and the functions $a_j(x)$ are of class C' on this interval. Moreover, on $x_1 t_g$ we have $a_j(x) = c_j$ $(j = 1, \cdots, n)$. In view of condition III′ and the Clebsch transformation of the second variation [I, p. 739], it would follow that

$$(3.3) \qquad \int_{x_1}^{x_3} 2\omega[x,\eta,\eta']dx + \eta_i(x_1)V_{ij}(x_1;\rho_1)c_j$$
$$= \int_{x_1}^{x_3} F_{y'_i y'_j}\{U_{ik}(x;\rho_1)a'_k(x)\}\{U_{jl}(x;\rho_1)a'_l(x)\}dx \geq 0.$$

On the other hand, by direct integration we obtain

$$\int_{x_1}^{x_3} 2\omega[x,\eta,\eta']dx = -\eta_i(x_1)\zeta_i(x_1),$$

and as a consequence of the initial values of $U_{ij}(x;\rho)$ and $V_{ij}(x;\rho)$, it would follow that

$$(3.4) \qquad \int_{x_1}^{x_3} 2\omega[x,\eta,\eta']dx + \eta_i(x_1)V_{ij}(x_1;\rho_1)c_j = -(\rho-\rho_1)\sum_{j=1}^{r_g} c_j^2.$$

Relation (3.4) is seen to be a contradiction to (3.3) unless $c_\tau = 0$ $(\tau = 1, \cdots, r_g)$. In this latter case, since $| U_{ij}(x_3;\rho_1) | \neq 0$ it would follow that $c_j = 0$ $(j = 1, \cdots, n)$, which is a contradiction. We have proved, therefore, that if $| U_{ij}(x;\rho_1) | \neq 0$ on $x_1 < x \leq t_\sigma + \epsilon$, then for $\rho > \rho_1$, we have $| U_{ij}(x;\rho) | \neq 0$ on this interval.

Finally, it is to be noted that on $t_\sigma + \epsilon \leq x \leq t_{\sigma-1}$ the determinant $| U_{ij}(x;\rho) |$ is a polynomial in ρ of degree $r_g - r_{\sigma-1}$ whose leading coefficient is $| u_{ij}(x;\sigma-1) |$, and therefore different from zero. Hence for ρ sufficiently large in absolute value we have $| U_{ij}(x;\rho) | \neq 0$ on $t_\sigma + \epsilon \leq x \leq t_{\sigma-1}$. Combining these results, we have that there exists a positive value ρ_2 such that $\rho_2 > \rho_1$ and $| U_{ij}(x;\rho_2) | \neq 0$ on $x_1 < x \leq t_{\sigma-1}$. We have established, therefore, an induction proof of Lemma 3.1.

In order to complete the proof of Theorem A, let $x_0 = x_1 - (t_g - x_1)$, and define the coefficients of ω and Φ_a on $x_0 \leq x < x_1$ by the following identities in (x, η, η'):

(3.5) $\quad \omega[x,\eta,\eta'] \equiv \omega[2x_1 - x, -\eta, \eta'], \quad \Phi_a[x,\eta,\eta'] \equiv \Phi_a[2x_1 - x, -\eta, \eta'].$

In the canonical system (2.3′) we then have for $x_0 \leq x < x_1$,

$$(3.6) \qquad A_{ij}(x) = -A_{ij}(2x_1 - x), \quad B_{ij}(x) = B_{ij}(2x_1 - x),$$
$$C_{ij}(x) = C_{ij}(2x_1 - x).$$

If $\eta_i(x) \equiv 0$, $\zeta_i(x)$ is a solution of (2.3′) on x_0x_1, then $\eta_i(x) \equiv 0$, $\zeta_i(2x_1 - x)$ is a solution of this system on x_1t_g, and conversely. Hence on every subinterval x_3x_1 of x_0x_1 the order of anormality is r_g. Moreover, on x_0x_1, the condition III′ holds.

It is to be emphasized that after the coefficients of the system (2.3), or (2.3′), have been defined on x_0x_1 by relations (3.5), these coefficients will in general be discontinuous at $x = x_1$. Hence, by a solution of this system on x_0x_2 we shall understand a set of functions $\eta_i(x)$, $\zeta_i(x)$ which are continuous on this interval, and which, except possibly at x_1, have derivatives satisfying equations (2.3′). We shall still denote by $u_{ij}(x)$, $v_{ij}(x)$, $x_0 \leq x \leq x_2$, the solutions of this modified system on x_0x_2 which satisfy the initial conditions (2.4). A set of r_g linearly independent solutions with $\eta_i(x) \equiv 0$ on x_0x_1 is given by the functions $u_{i\tau}(x)$, $v_{i\tau}(x)$ $(\tau = 1, \cdots, r_g)$.

Now let $v^*_{ij}(x)$ $(j = 1, \cdots, n)$ be continuous functions such that $v^*_{ij}(x_1) = v_{ij}(x_1)$, and

$$(3.7) \qquad v_{i\tau}(x)v^*_{ij}(x) \equiv 0 \text{ on } x_0x_1 \text{ if } \tau \neq j \ (\tau = 1, \cdots, r_g;\ j = 1, \cdots, n).$$

Moreover, let ρ be such that $|U_{ij}(x;\rho)| \neq 0$ on $x_1 < x \leq x_2$.

Corresponding to a point t on x_0x_1 there exist mutually conjugate solutions $\eta_{ij}(x;t)$, $\zeta_{ij}(x;t)$ of (2.3′) on x_0x_2 such that

$$(3.8) \qquad \begin{array}{lll} \eta_{ij}(t;t) = v_{ij}(t), & \zeta_{ij}(t;t) = \rho v^*_{ij}(t) & \text{for } (j = 1, \cdots, r_g), \\ \eta_{ij}(t;t) = 0, & \zeta_{ij}(t;t) = v^*_{ij}(t) & \text{for } (j = r_g + 1, \cdots, n). \end{array}$$

For $t = x_1$ the initial conditions (3.8) reduce to the initial values of $U_{ij}(x;\rho)$, $V_{ij}(x;\rho)$. Hence, $\eta_{ij}(x;x_1) \equiv U_{ij}(x;\rho)$, $\zeta_{ij}(x;x_1) \equiv V_{ij}(x;\rho)$ on x_0x_2.

It will now be shown that for $t < x_1$ and sufficiently near to x_1 the determinant $|\eta_{ij}(x;t)| \neq 0$ on x_1x_2. It will first be noted that on a subinterval $x'x''$, where $x_0 \leq x' < x_1$ and $x_1 \leq x'' \leq t_g$, the order of anormality is r_g. The following lemma may then be proved by the method used by Bliss [I, p. 739] to prove a corresponding result for an arc which satisfies stronger normality conditions.

Lemma 3.2. *There exists an interval* $x_1 - d \leq x \leq x_1 + d$ $(0 < d < t_g - x_1)$ *such that if* $\eta_i(x)$, $\zeta_i(x)$ *is a solution of* (2.3′), *and the functions* $\eta_i(x)$ *all vanish at a point* x' *of* $x_1 - d \leq x < x_1$ *and at a*

point x'' of $x_1 \leq x \leq x_1 + d$, then on $x'x''$ we have $\eta_i(x) \equiv 0$, $\zeta_i(x) = v_{i\tau}(x)c_\tau$, where c_τ $(\tau = 1, \cdots, r_g)$ are constants.

From the initial values of the functions $v^*_{ij}(x)$ it is seen that there exists a δ such that $0 < \delta < d$, where d is determined as in Lemma 3.2, and such that if $x_1 - \delta \leq x \leq x_1$, then the determinant $| v_{i\tau}(x) v^*_{i\sigma}(x) |$ $(\tau = 1, \cdots, r_g;\ \sigma = r_g + 1, \cdots, n)$ is different from zero. It then follows from the result of Lemma 3.2, that if t is an arbitrary point on $x_1 - \delta \leq t < x_1$ and $\eta_{ij}(x; t)$, $\zeta_{ij}(x; t)$ is the corresponding family determined by the initial conditions (3.8), then $| \eta_{ij}(x; t) | \neq 0$ on $x_1 \leq x \leq x_1 + d$.

Finally, since $\eta_{ij}(x; x_1) \equiv U_{ij}(x; \rho)$, it follows from the continuity of the solutions $\eta_{ij}(x; t)$, $\zeta_{ij}(x; t)$ of (2.3′) when considered as functions of t, that for t sufficiently near to x_1 the determinant $| \eta_{ij}(x; t) |$ is also different from zero on $x_1 + d \leq x \leq x_2$. Therefore, for $t < x_1$ and sufficiently near to x_1 the family of mutually conjugate solutions of (2.3′) determined by the conditions (3.8) is such that $| \eta_{ij}(x; t) | \neq 0$ on $x_1 \leq x \leq x_2$. This completes the proof of Theorem A.

The following corollary is an immediate consequence of well-known results for the problem of Lagrange:

COROLLARY. *Suppose E_{12} is an extremal arc which satisfies the conditions of Theorem A. If $u_i(x)$, $v_i(x)$ is an accessory extremal, and $\eta_i(x)$ is an arbitrary admissible variation such that $\eta_i(x_1) = u_i(x_1)$, $\eta_i(x_2) = u_i(x_2)$, then*

$$\int_{x_1}^{x_2} 2\omega[x, \eta, \eta']\,dx \geq \int_{x_1}^{x_2} 2\omega[x, u, u']\,dx, \tag{3.9}$$

and the equality sign holds if and only if $\eta_i(x) \equiv u_i(x)$ on x_1x_2.

4. Further discussion of Theorem A. It has been established in Lemma 3.1 that along an extremal arc which satisfies the conditions of Theorem A the accessory extremals $U_{ij}(x; \rho)$, $V_{ij}(x; \rho)$ defined by (3.1) are such that $| U_{ij}(x; \rho) | \neq 0$ on $x_1 < x \leq x_2$ for $\rho > 0$ and sufficiently large in value. If $r_g = r_0$, that is, if the order of anormality of E_{12} is the same on every sub-interval x_1x of x_1x_2, it is seen from the first paragraph of the proof of Lemma 3.1 that this condition is true for arbitrary values of ρ. For further discussion we shall assume $r_g > r_0$, and seek to determine a lower bound for the values of ρ which are such that the conclusion of Lemma 3.1 is satisfied. This lower bound is not determined independent of the results of § 3, however, since use is made of the Corollary of that section.

In view of condition IV′$_0$ and condition (α) of property 5° in § 2, it is

seen that the $2n - r_0$ accessory extremals $u_{i\kappa}(x)$, $v_{i\kappa}(x)$ $(\kappa = r_0 + 1, \cdots, 2n)$ defined by (2. 4) and (2. 5) are such that the matrix

$$\left\| \begin{matrix} u_{i\kappa}(x_1) \\ u_{i\kappa}(x_2) \end{matrix} \right\|$$

is of rank $2n - r_0$. As a consequence of property 2° of § 2, or by a direct consideration of the initial values $u_{is}(x_1)$ $(s = 1, \cdots, 2n)$, it is seen that the matrix of $2n - r_g$ rows and $2n - r_0$ columns

$$\left\| \begin{matrix} v_{i\sigma}(x_1) u_{i\kappa}(x_1) \\ u_{i\kappa}(x_2) \end{matrix} \right\| \qquad \begin{matrix} (\sigma = r_g + 1, \cdots, n; \; i = 1, \cdots, n; \\ \kappa = r_0 + 1, \cdots, 2n) \end{matrix} \tag{4.1}$$

is of rank $2n - r_g$.

For simplicity of notation, let $2A[z] \equiv a_{\kappa\lambda} z_\kappa z_\lambda$ $(\kappa, \lambda = r_0 + 1, \cdots, 2n)$ denote the quadratic form

$$\int_{x_1}^{x_2} 2\omega[x, u_{i\kappa} z_\kappa, u'_{i\kappa} z_\kappa] dx = z_\kappa u_{i\kappa}(x) v_{i\lambda}(x) z_\lambda \,\Big|_{x=x_1}^{x=x_2} . \tag{4.2}$$

We shall also denote by $B_\nu[z] \equiv b_{\nu\kappa} z_\kappa$ $(\nu = 1, \cdots, 2n - r_g)$ the first members of the linear equations

$$\begin{matrix} v_{i\sigma}(x_1) u_{i\kappa}(x_1) z_\kappa = 0, \\ u_{i\kappa}(x_2) z_\kappa = 0, \end{matrix} \qquad \begin{matrix} (\sigma = r_g + 1, \cdots, n; \; i = 1, \cdots, n; \\ \kappa = r_0 + 1, \cdots, 2n) \end{matrix} \tag{4.3}$$

Finally, denote by $2K[z] \equiv k_{\kappa\lambda} z_\kappa z_\lambda$ the quadratic form

$$[u_{i\kappa}(x_1) u_{i\lambda}(x_1) + u_{i\kappa}(x_2) u_{i\lambda}(x_2)] z_\kappa z_\lambda, \tag{4.4}$$

which, in view of IV'_0 and elementary properties of matrices, is positive definite. It then follows that the class of values (z_κ) which satisfy the conditions

$$B_\nu[z] = 0, \qquad 2K[z] = 1 \tag{4.5}$$

is not vacuous. Moreover, the minimum value of $2A[z]$ in this class of values is the smallest zero l_1 of the determinant [9]

---

[9] See, for example, Hancock, *Theory of Maxima and Minima*, Ginn and Co., Boston (1917), pp. 103-114. Bliss has phrased his analogue of the Jacobi condition for the problem of Bolza in terms of the roots of a determinant of the form (4. 6); see [IV], p. 273. It may be shown that corresponding to each of the zeros $l = l_h$ $(h = 1, \cdots, r_g - r_0)$ of $D(l)$ there exists an accessory extremal $\eta_{ih}(x)$, $\zeta_{ih}(x)$ satisfying with constants $d_{\sigma h}$ the end conditions:

$$v_{i\sigma}(x_1)\eta_{ih}(x_1) = 0, \; \eta_{ih}(x_2) = 0; \; d_{\sigma h} v_{i\sigma}(x_1) - l_h \eta_{ih}(x_1) - \zeta_{ih}(x_1) = 0.$$

See [IX], § 5, and also [V].

$$D(l) = \begin{vmatrix} a_{\kappa\lambda} - lk_{\kappa\lambda} & b_{\nu\kappa} \\ b_{\nu\lambda} & 0 \end{vmatrix}. \tag{4.6}$$

Finally, in view of IV$'_0$ we have that if $\eta_i(x)$ is an arbitrary admissible variation, then there exists a unique accessory extremal $u_i = u_{i\kappa}(x)z_\kappa$, $v_i = v_{i\kappa}(x)z_\kappa$, such that $u_i(x_1) = \eta_i(x_1)$, $u_i(x_2) = \eta_i(x_2)$ [X, p. 809]. As a consequence of the minimizing property of l_1 and the corollary of § 3, we have that if $\eta_i(x)$ is an admissible variation such that

$$v_{i\sigma}(x_1)\eta_i(x_1) = 0, \quad \eta_i(x_2) = 0, \quad \eta_i(x_1)\eta_i(x_1) \neq 0, \tag{4.7}$$

then
$$\int_{x_1}^{x_2} 2\omega[x, \eta, \eta']dx \geq l_1[\eta_i(x_1)\eta_i(x_1)].$$

Now suppose that for a given value of ρ there exists a point x_3 such that $x_1 < x_3 \leq x_2$ and $|U_{ij}(x_3; \rho)| = 0$. If c_j $(j = 1, \cdots, n)$ are constants not all zero and such that $U_{ij}(x_3; \rho)c_j = 0$, it follows from IV$'_0$ that $c_\tau c_\tau \neq 0$, $(\tau = 1, \cdots, r_g)$. Moreover, the admissible arc defined by

$$\eta_i(x) = U_{ij}(x; \rho)c_j \text{ on } x_1x_3; \quad \eta_i(x) \equiv 0 \text{ on } x_3x_2,$$

satisfies equations (4.7). On direct integration, we obtain

$$\int_{x_1}^{x_2} 2\omega[x, \eta, \eta']dx = -\rho c_\tau c_\tau = -\rho[\eta_i(x_1)\eta_i(x_1)],$$

and we must have, therefore, $\rho \leq -l_1$. We have established, therefore, the following theorem:

THEOREM 4.1. *Suppose that E_{12} is an extremal arc satisfying the conditions of Theorem A, and $U_{ij}(x; \rho)$, $V_{ij}(x; \rho)$ is the system of accessory extremals defined by* (3.1). *Then for $\rho > -l_1$, where l_1 is the smallest zero of the determinant $D(l)$, we have $|U_{ij}(x; \rho)| \neq 0$ on $x_1 < x \leq x_2$.*

Finally, there will be given a method for determining a system of mutually conjugate accessory extremals $\eta_{ij}(x)$, $\zeta_{ij}(x)$ with $|\eta_{ij}(x)| \neq 0$ on x_1x_2 which does not use directly the system defined by (3.1).

Consider the problem of minimizing $2A[z]$ in the class of values (z_κ) $(\kappa = r_0 + 1, \cdots, 2n)$ which satisfy the conditions

$$u_{i\kappa}(x_2)\, z_\kappa = 0, \qquad 2K[z] = 1, \tag{4.8}$$

where $A[z]$ and $K[z]$ are defined by (4.2) and (4.4). The minimum value of $2A[z]$ in this class of values is the smallest zero m_1 of the determinant

$$(4.9) \qquad \mathfrak{D}(m) = \begin{vmatrix} a_{\kappa\lambda} - m k_{\kappa\lambda} & u_{j\kappa}(x_2) \\ u_{i\lambda}(x_2) & 0 \end{vmatrix}.$$

By the use of the Corollary of § 3, and by an argument like that used above, it is seen that if $\eta_i(x)$ is an admissible variation such that $\eta_i(x_2) = 0$, $\eta_i(x_1)\eta_i(x_1) \neq 0$, then

$$\int_{x_1}^{x_2} 2\omega[x, \eta, \eta']dx \geq m_1[\eta_i(x_1)\eta_i(x_1)].$$

In terms of the accessory extremals (2.4) and (2.5), now define a system of mutually conjugate accessory extremals $Y_{ij}(x;\rho)$, $Z_{ij}(x;\rho)$ as follows

$$(4.10) \qquad Y_{ij}(x;\rho) = u_{i\,n+j}(x) + \rho u_{ij}(x), \quad Z_{ij}(x;\rho) = v_{i\,n+j}(x) + \rho v_{ij}(x).$$

Suppose that for a given value of ρ there exists a point x_3 such that $x_1 < x_3 \leq x_2$ and $| Y_{ij}(x_3;\rho)| = 0$. Let d_j $(j = 1, \cdots, n)$ be constants not all zero such that $Y_{ij}(x_3;\rho)d_j = 0$. If we define

$$\eta_i(x) = Y_{ij}(x;\rho)d_j \text{ on } x_1x_3, \quad \eta_i(x) \equiv 0 \text{ on } x_3x_2,$$

then $\eta_i(x_1)\eta_i(x_1) = d_id_i \neq 0$. Moreover, by direct integration we obtain

$$\int_{x_1}^{x_2} 2\omega[x, \eta, \eta']dx = -\rho d_id_i = -\rho[\eta_i(x_1)\eta_i(x_1)].$$

Hence, if ρ is a value such that $| Y_{ij}(x;\rho)| = 0$ at a point x_3 on $x_1 < x \leq x_2$, it follows that $\rho \leq - m_1$. Finally, it is to be noted that

$$| Y_{ij}(x_1;\rho)| = | u_{i\,n+j}(x_1)| \neq 0.$$

We have, therefore, the following theorem.

THEOREM 4.2. *Suppose that E_{12} is an extremal arc satisfying the conditions of Theorem A, and that $Y_{ij}(x;\rho)$, $Z_{ij}(x;\rho)$ is the family of mutually conjugate accessory extremals defined by* (4.10). *Then for $\rho > - m_1$, where m_1 is the smallest zero of the determinant $\mathfrak{D}(m)$, we have $| Y_{ij}(x;\rho)| \neq 0$ on $x_1 \leq x \leq x_2$.*

In Theorem 4.2 we have assumed that E_{12} is an extremal satisfying the conditions of Theorem A, that is, it satisfies the strengthened Clebsch condition and has on it no point conjugate to the point 1. Consequently, in the proof of Theorem 4.2 use is made of the Corollary of § 3. It is significant to note, however, that quite independent of the results of § 3 the above proof of Theorem 4.2 leads to a result which is of itself important.

In the proof of the above theorem explicit use is made of only the following hypotheses: (1) *E_{12} is an extremal arc satisfying the strengthened Clebsch condition,* (2) *the point 2 is not conjugate to the point* 1, *and* (3) *if $u_i(x)$, $v_i(x)$ is an accessory extremal and $\eta_i(x)$ is an arbitrary admissible variation such that $\eta_i(x_1) = u_i(x_1)$, $\eta_i(x_2) = u_i(x_2)$, then the inequality* (3.9) *holds.* Let μ_a denote the multipliers corresponding to the accessory extremal u_i, v_i. Then by an expansion of the integrand function $2\omega[x, \eta, \eta'] = 2\Omega[x, \eta, \eta', \mu]$ by Taylor's formula, and an integration by parts, one obtains that assumption (3) is equivalent to the condition that if $\eta_i(x)$ is an arbitrary admissible variation such that $\eta_i(x_1) = 0 = \eta_i(x_2)$, then

$$(4.10) \qquad \int_{x_1}^{x_2} 2\omega[x, \eta, \eta']dx \geq 0.$$

In view of this remark we have, therefore, that quite independent of § 3 the above proof of Theorem 4.2 leads to the following result:

THEOREM 4.3. *Suppose E_{12} is an extremal arc which satisfies the strengthened Clebsch condition, on which the point* 2 *is not conjugate to the point* 1, *and such that if $\eta_i(x)$ is an arbitrary admissible variation along E_{12} having $\eta_i(x_1) = 0 = \eta_i(x_2)$, then inequality* (4.10) *is satisfied. Then for $\rho > -m_1$, where m_1 is the smallest zero of the determinant $\mathfrak{D}(m)$, we have $|Y_{ij}(x;\rho)| \neq 0$ on $x_1 \leq x \leq x_2$.*

It is to be remarked that one may show the equivalence of the hypotheses of Theorems 4.2 and 4.3 by the use of the results of § 3 and the first necessary condition for the problem of Lagrange. However, the significance of Theorem 4.3 lies in its simplicity of proof and its usefulness. For example, with the aid of this theorem one may establish in a simple manner the existence of infinitely many characteristic numbers for the boundary value problem which the author has previously treated by the use of other methods [see V and IX].

5. Relation of Theorem A to sufficiency theorems for the problem of Bolza. By the use of Theorem A the proof of a sufficiency theorem for the problem of Bolza as given by Bliss [IV] is still valid when the normality conditions there used are replaced by the weaker assumptions of normality with respect to the end conditions and normality on x_1x_2. This latter assumption of normality is used by Bliss not only in the proof of the imbedding theorem, but also in the proof of another result [IV, Lemma 2, p. 267]. Hence, Theorem A does not eliminate entirely the assumption of normality

on x_1x_2 in the proof of sufficient conditions as given by Bliss. The result of Theorem A also enables one to establish sufficiency theorems for the problem of Mayer by the methods used by Bliss and Hestenes [VI] and Hestenes [VII] under correspondingly weakened hypotheses.

In a recent paper [IX] the author has discussed for the problem of Mayer the relations between the boundary value problem formulation of the analogue of Jacobi's condition and the analogue of this condition introduced by Bliss {see [IV] and [VII]}. As a consequence of Theorem A the results of Theorems 4. 1, 4. 2 and 4. 4 of that paper are valid when the assumptions concerning normality on sub-intervals is omitted. Consequently, these results apply directly to the problem of Bolza as well as to the problem of Mayer. In particular, using the terminology of that paper, we have:

If E_{12} is an extremal of the form $y_i = y_i(x)$, $\lambda_0 = 1$, $\lambda_\alpha(x)$ which satisfies conditions I and III′, then condition IV′$_C$ is satisfied if and only if conditions IV′$_0$ and IV′$_B$ are satisfied.

For the problem of Bolza as treated by Hestenes [X], Theorem A leads to the following results, when expressed in Hestenes' notation [X, pp. 810-811].

If g is an extremal satisfying conditions I and III′, then condition V′ is satisfied if and only if condition VI′ is satisfied along g. Again, using the notation of Hestenes, we have the following relation between the conditions IV′$_C$, IV′$_0$ and IV′$_B$ of [IX], and conditions V′ and VI′ of Hestenes {see [X], pp. 810-816}.

Suppose that the end conditions are regular, and that the non-tangency condition holds on an admissible arc g having no corners. If g satisfies conditions I, III′ with a set of multipliers $\lambda_0 = 1$, $\lambda_\beta(x)$, then each of the four following conditions implies the others: IV′$_C$, IV′$_0$ with IV′$_B$, V′, VI′. In case the end conditions are separated, then each of the above conditions is equivalent to the condition IV′ of Hestenes [X, p. 806].

The above results lead to obvious simplification of the sufficiency theorems as stated by Hestenes.

BIBLIOGRAPHY.

I. Bliss, "The problem of Lagrange in the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 673-747.

II. Morse, "Sufficient conditions in the problem of Lagrange with fixed end points," *Annals of Mathematics*, ser. 2, vol. 32 (1931), pp. 567-577.

III. Morse, "Sufficient conditions in the problem of Lagrange with variable end conditions," *American Journal of Mathematics*, vol. 53 (1931), pp. 517-546.

IV. Bliss, "The problem of Bolza in the calculus of variations," *Annals of Mathematics*, ser. 2, vol. 33 (1932), pp. 261-274.

V. Reid, "A boundary value problem associated with the calculus of variations," *American Journal of Mathematics*, vol. 54 (1932), pp. 769-790.

VI. Bliss and Hestenes, "Sufficient conditions for a problem of Mayer in the calculus of variations," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 305-326; *Contributions to the Calculus of Variations* 1931-1932, The University of Chicago Press, pp. 295-337.

VII. Hestenes, "Sufficient conditions for the general problem of Mayer with variable end points," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 479-490; *Contributions to the Calculus of Variations* 1931-1932, The University of Chicago Press, pp. 339-360.

VIII. Hu, "The problem of Bolza and its accessory boundary value problem," *Contributions to the Calculus of Variations* 1931-1932, The University of Chicago Press, pp. 361-443.

IX. Reid, "Analogues of the Jacobi condition for the problem of Mayer in the calculus of variations," *Annals of Mathematics*, ser. 2, vol. 35 (1934), pp. 836-848.

X. Hestenes, "Sufficient conditions for the problem of Bolza in the calculus of variations," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 793-818.

The University of Chicago,
Chicago, Illinois.

CONCERNING SOME METHODS OF BEST APPROXIMATION, AND A THEOREM OF BIRKHOFF.[1]

By I. M. Sheffer.

Introduction. The series of Taylor has been generalized in many directions. It is our purpose here to consider a natural extension of certain series in addition to that of Taylor, by methods analogous to those used by Birkhoff[2] and by Widder[3] for Taylor series.

The Widder theory is based on a best approximation definition. Let $\{\phi_n(x)\}$, $(n=0,1,\cdots)$ be an infinite sequence of functions, analytic about $x=0$. An expression of the form $s_n(x)=\sum_0^n c_i\phi_i(x)$ is a "polynomial" of order (not exceeding) n. According to Widder, the "polynomial" $s_n(x)$ is the best approximation of order n to the function $f(x)$, analytic about $x=0$, if $f(x)-s_n(x)$ vanishes, together with its first n derivatives, at $x=0$. If certain determinants do not vanish, then s_n exists and is unique. The question of convergence of $s_n(x)$ to $f(x)$ is equivalent to that of $s_0(x)+\sum_1^\infty[s_n(x)-s_{n-1}(x)]$ to $f(x)$. Now a striking situation comes to light: s_n-s_{n-1} can be written in the form

$$s_n(x)-s_{n-1}(x)=f_n\Omega_n(x)$$

where $\Omega_n(x)$ is a "polynomial" of order n, *independent of* $f(x)$, the function $f(x)$ making its presence felt only in the constants $\{f_n\}$. That is, $s_0(x)+\sum_1^\infty[s_n(x)-s_{n-1}(x)]\sim\sum_0^\infty f_n\Omega_n(x)$; so that in going from the n-th approximation to the $(n+1)$-st, we have only to add in the term $f_{n+1}\Omega_{n+1}(x)$, the terms already present remaining. Such behavior (which is also true of the extension we have in view) we shall refer to as the *property of permanence.*

In the Widder case,

$$\Omega_n(x)=(x^n/n!)[1+h_n(x)],$$

[1] Presented to the American Mathematical Society, December, 1933.

[2] G. D. Birkhoff, "Sur une généralisation de la série de Taylor," *Comptes Rendus*, vol. 164 (1917), pp. 942-945.

[3] D. V. Widder, "On the expansion of analytic functions of the complex variable in generalized Taylor's series," *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 43-52. (This paper contains a reference to a preceding work by Widder in which functions of a *real variable* are considered; but that phase of the problem does not concern us here.)

where $h_n(x)$ is analytic at $x = 0$ and $h_n(0) = 0$; and his convergence theory is based on the following additional hypotheses:

(i) $h_n(x)$ is analytic, $| x | \leqq R$;

(ii) M exists, independent of n and x, such that $| h_n(x) | \leqq M/(n+1)$, $| x | \leqq R$.

In § 1 we give a rather general definition of best approximation, and establish the property of permanence. As particular cases are the Widder case, and the "least square" case. In § 2 we show that condition (ii) of Widder can be replaced by the less restrictive condition:

(ii′) $1 + h_n(x)$ converges uniformly, in $| x | \leqq R$, to a function $M(x)$, with $M(x) \neq 0$ in $| x | \leqq R$.

Essentially, we replace Widder's condition $| h_n(x) | = 0(1/n)$ by the condition $| h_n(x) | = o(1)$. Finally, in § 3, to handle convergence in some other cases, we appeal to the method used by Birkhoff, which by means of an integral equation extends the convergence properties of Taylor series (i. e., series in $\{x^n\}$) to series in $\{v_n(x)\}$, where $\{v_n(x)\}$ is "sufficiently close" to $\{x^n\}$. Only, the rôle of $\{x^n\}$ is now played by functions $\{u_n(x)\}$, which we endow with properties analogous to those of the known functions $\{x^n\}$.

1. *Some methods of best approximation.* Let $\{\phi_n(x)\}$, $(n = 0, 1, \cdots)$ be a sequence of functions. We wish to assign to each function $f(x)$ (of a certain class) a sequence of "polynomials" $\{s_n(x)\}$:

$$s_n(x) = c_{n0}\phi_0(x) + \cdots + c_{nn}\phi_n(x), \tag{1}$$

which are the *best* approximating "polynomials" to $f(x)$, each of its order. This requires that we define the test for best approximation.

Let L_n, $(n = 0, 1, \cdots)$ be a sequence of linear operators, each of which assigns to a function $u(x)$ a number: $L_n[u(x)] = u_n$.

Definition. *By the method $\mathfrak{M}$ of best approximation, relative to the set of operators L_n, is meant that determination of the set $\{s_n(x)\}$ according to the following test of best approximation:*[4]

$$L_i[s_n(x)] = L_i[f(x)], \qquad (i = 0, 1, \cdots, n). \tag{2}$$

[4] The "polynomials" $s_n(x)$ depend, of course, on the sequence of functions $\{\phi_n(x)\}$.

DEFINITION. *A method $\mathfrak{M}$ is non-singular, relative to a sequence $\{\phi_n(x)\}$, if none of the following determinants vanishes:* [5]

$$(3) \qquad \Delta_n = \begin{vmatrix} L_0[\phi_0] & L_0[\phi_1] & \cdots & L_0[\phi_n] \\ L_1[\phi_0] & L_1[\phi_1] & \cdots & L_1[\phi_n] \\ \cdot & \cdot & \cdot & \cdot \\ L_n[\phi_0] & L_n[\phi_1] & \cdots & L_n[\phi_n] \end{vmatrix}, \qquad (n = 0, 1, \cdots).$$

We shall consider only non-singular methods.

THEOREM 1. *For a given function* [6] *$f(x)$, to each $(n = 0, 1, \cdots)$ there is a unique best approximating "polynomial" $s_n(x)$ of order (not greater than) n.*

This follows from equations (2), since $s_n(x)$ has the form (1), so that the determinant of system (2) is precisely $\Delta_n \neq 0$.

Let us form from $\{\phi_n(x)\}$ a new sequence $\{\Phi_n(x)\}$, linearly dependent on $\{\phi_n(x)\}$:

$$(4) \qquad \Phi_n(x) = b_{n0}\phi_0(x) + \cdots + b_{nn}\phi_n(x), \qquad b_{nn} \neq 0.$$

Clearly,[7] the set of best approximating "polynomials" for $\{\Phi_n(x)\}$ coincides with the set $\{s_n(x)\}$ already found for $\{\phi_n(x)\}$. We may therefore choose one set out of the infinite number of possible ones, to be obtained from $\{\phi_n(x)\}$ as was $\{\Phi_n(x)\}$, to represent all such. There exists a "significant" set, which we shall term the basic set.

DEFINITION. *The basic set for a given sequence $\{\phi_n(x)\}$, relative to a method $\mathfrak{M}$, is the set $\{\Phi_n(x)\}$ defined by*

$$(5) \qquad L_i[\Phi_n(x)] = \begin{cases} 0, & (i = 0, 1, \cdots, n-1); \\ 1, & (i = n), \end{cases}$$

where $\Phi_n(x)$ has the form (4).

In virtue of the condition $\Delta_n \neq 0$, we see that [8] *a basic set exists and is unique.*

---

[5] A method $\mathfrak{M}$ may be essentially singular; i.e., is singular for *all* sequences $\{\phi_n(x)\}$, as for example if n exists such that $L_0, L_1, \cdots, L_n$ are linearly dependent. On the other hand, $\mathfrak{M}$ may be in general non-singular, but a peculiar choice of $\{\phi_n(x)\}$ may give singularity, as for example if n exists such that $\phi_0, \cdots, \phi_n$ are linearly dependent.

[6] It is understood that $f(x)$ and the functions $\phi_n(x)$ are within the class of functions on which the L_n's can operate.

[7] The determinants Δ_n for $\{\Phi_n(x)\}$ are non-vanishing, as is easily seen from (3).

[8] The only point not obvious is that in $\Phi_n(x)$, $b_{nn} \neq 0$. But if $b_{nn} = 0$, then the

LEMMA 1. *There is a sequence of constants* $\{f_n\}$ such that

$$(6) \quad s_0(x) = f_0\Phi_0(x); \qquad s_n(x) - s_{n-1}(x) = f_n\Phi_n(x), \quad (n = 1, 2, \cdots).$$

Since s_0 and Φ_0 are each multiples of ϕ_0, and $b_{00} \neq 0$, therefore f_0 can be found. Now consider $T_n = s_n - s_{n-1}$. We have from (2):

$$L_i[T_n] = 0, \quad (i = 0, 1, \cdots, n-1); \qquad L_n[T_n] = f_n,$$

where we set

$$(7) \qquad f_n = L_n[s_n(x) - s_{n-1}(x)].$$

If $f_n = 0$, then since $\Delta_n \neq 0$ we have $T_n(x) \equiv 0$, thus satisfying (6). If $f_n \neq 0$, then T_n/f_n satisfies equations (5), whence by uniqueness, $T_n/f_n \equiv \Phi_n$, and again (6) holds.

COROLLARY 1. *Method* $\mathfrak{M}$ *has the permanence property:*[9]

$$(8) \qquad s_n(x) = s_0(x) + \sum_{i=1}^{n} [s_i(x) - s_{i-1}(x)] = \sum_{i=0}^{n} f_i\Phi_i(x),$$

where only the constants f_i *(which are independent of* n*) depend on the function* $f(x)$.

COROLLARY 2. *The constants* $\{f_n\}$ *are given by*

$$(9) \qquad f_n = \frac{1}{\Delta_{n-1}} \begin{vmatrix} L_0[\phi_0] & L_0[\phi_1] & \cdots & L_0[\phi_{n-1}] & L_0[f] \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ L_n[\phi_0] & L_n[\phi_1] & \cdots & L_n[\phi_{n-1}] & L_n[f] \end{vmatrix},$$

$$(n = 1, 2, \cdots); \qquad f_0 = L_0[f].$$

For, let g_n denote the determinant in the numerator of the right-hand side of (9). In (6), $s_n - s_{n-1}$ and $f_n\Phi_n$ are linear combinations of the functions $\phi_0, \cdots, \phi_n$ which are (as we have observed in a footnote) linearly independent. Hence coefficients of ϕ_n on both sides of (6) must be equal. From (2) this coefficient in $s_n - s_{n-1}$ is g_n/Δ_n, and from (5) this coefficient in $f_n\Phi_n$ is $f_n\Delta_{n-1}/\Delta_n$. Hence $f_n = g_n/\Delta_{n-1}$, which is (9).

COROLLARY 3. *The functions* $\{\Phi_n(x)\}$ *are given by*

---

first n equations of (5) tell us, since $\Delta_{n-1} \neq 0$, that $b_{n0} = b_{n1} = \cdots = b_{n,n-1} = 0$, so that the $(n+1)$-st equation of (5) is not satisfied; a contradiction.

[9] From this follows the curious fact that if we choose $f(x) = \Phi_n(x)$, then *zero* is the best approximating "polynomial" of all orders less than n.

$$(10)\qquad \Phi_n(x) = \frac{1}{\Delta_n}\begin{vmatrix} L_0[\phi_0] \cdots & L_0[\phi_n] \\ \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ L_{n-1}[\phi_0] \cdots & L_{n-1}[\phi_n] \\ \phi_0(x) \cdots & \phi_n(x) \end{vmatrix},$$

$$(n = 1, 2, \cdots); \qquad \Phi_0(x) = \phi_0(x)/\Delta_0.$$

(10) follows at once from (5) and the uniqueness of a basic set.

COROLLARY 4. *The "polynomials" $s_n(x)$ are given by*

$$(11)\quad s_n(x) = \frac{-1}{\Delta_n}\begin{vmatrix} L_0[\phi_0] \cdots L_0[\phi_n] & L_0[f] \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \\ L_n[\phi_0] \cdots L_n[\phi_n] & L_n[f] \\ \phi_0(x) \cdots \phi_n(x) & 0 \end{vmatrix}, \qquad (n = 0, 1, 2, \cdots).$$

For, operate on the right-hand member of (11) with L_i $(0 \leqq i \leqq n)$, and subtract the resulting last row from the row with index i, using the property $L_i[0] = 0$. Then expand in terms of the elements in this row of index i. There results $L_i[f]$, $(i = 0, 1, \cdots, n)$. That is, the right side of (11) is a linear combination of $\phi_0, \cdots, \phi_n$ satisfying the conditions (2). But system (2) has a unique solution $s_n(x)$; whence (11) follows.

Let us now return to the definition of a best approximation method. If we start with a set of linear operators M_n, which assign *functions* to functions: $M_n[u(x)] = u_n(x)$, then by choosing a sequence of numbers $\{x_n\}$, we get a method $\mathfrak{M}$ by setting $L_n[u(x)] = \{M_n[u(x)]\}_{x=x_n}$. In particular, we may have $x_n \equiv a$.

An interesting subclass is that where the operators M_n are obtained by iteration from a single one:

$$M_0 \equiv I \equiv \text{identity},\ M_1 = M,\ M_2 = M(M) = M^2, \cdots, M_n = M^n, \cdots;$$

and where we choose $x_n \equiv a$ (which we may take as 0). The case of Widder finds itself in this class, with $M[u(x)] \equiv du(x)/dx$.

For any method $\mathfrak{M}$ which is non-singular relative to the set $\{\phi_n(x) = x^n\}$, there will exist a unique basic set of *polynomials* $\Phi_n(x) = \sum_{i=0}^{n} \lambda_{n_i} x^i$. If $M[u] = du(x)/dx$, then $\Phi_n(x) = x^n/n!$. Again, if $M[u] = u(x+1) - u(x)$, then $\{\Phi_n(x)\}$ is the set of Newton polynomials:

$$\Phi_0(x) = 1, \quad \Phi_1(x) = x, \quad \Phi_n(x) = \frac{x(x-1)\cdots(x-n+1)}{n!}.$$

And in general, by means of these "methods" we can define large classes of sets of polynomials.

In particular, consider the class of orthogonal Tchebycheff polynomial sets. Given a function $p(x)$, a Tchebycheff set $\{T_n(x)\}$ is defined by the relations

$$\int_a^b p(t)T_m(t)T_n(t)\,dt = \begin{cases} 0, & (m \neq n), \\ 1, & (m = n). \end{cases}$$

These are equivalent (except for an undetermined multiplier in $T_n(x)$) to

$$\int_a^b p(t)t^iT_n(t)\,dt = 0, \qquad (i = 0, 1, \cdots, n-1).$$

If we now define

$$M_i[u(x)] = \int_a^b p(t)(t-x)^i u(t)\,dt, \qquad (i = 0, 1, \cdots),$$

then

$$L_i[T_n(x)] = \{M_i[T_n(x)]\}_{x=0} = 0, \qquad (i = 0, 1, \cdots, n-1),$$

and, by properly normalising $T_n(x)$, $L_n[T_n(x)] = 1$; so that $T_n(x) = \Phi_n(x)$. We thus see that *all orthogonal Tchebycheff polynomial sets can be defined by our methods of best approximation.*

More generally, the same is true for least square functions: Given the linearly independent set of functions $\{\phi_n(x)\}$, the set $\{T_n(x)\}$ is to be defined by

$$\int_a^b [f(t) - T_n(t)]^2\,dt = \text{minimum},$$

where $T_n(t)$ ranges over all "polynomials" $T_n(x) = c_{n0}\phi_0(x) + \cdots + c_{nn}\phi_n(x)$. By forming suitable linear combinations $\Phi_n(x) = b_{n0}\phi_0(x) + \cdots + b_{nn}\phi_n(x)$, $b_{nn} \neq 0$, we can make $\{\Phi_n\}$ a normal orthogonal set:

$$\int_a^b \Phi_m(t)\Phi_n(t)\,dt = \begin{cases} 0, & (m \neq n), \\ 1, & (m = n). \end{cases}$$

Furthermore, a method $\mathfrak{M}$ can be found for which $\{\Phi_n\}$ is the basic set. For, we have only to define

$$L_i[u(x)] = \int_a^b \Phi_i(t)u(t)\,dt;$$

then

$$L_i[\Phi_n(x)] = \begin{cases} 0, & (i = 0, 1, \cdots, n-1), \\ 1, & (i = n), \end{cases}$$

so that $\{\Phi_n\}$ *is* a basic set. Now we can express $T_n(x)$ as a linear combination

in the Φ_n's: $T_n(x) = f_{n0}\Phi_0 + \cdots + f_{nn}\Phi_n$. From the minimum property we find that

$$f_{ni} = f_i = \int_a^b f(t)\Phi_i(t)\,dt.$$

Again, $$L_i[T_n] = f_{ni}L_i[\Phi_i] = f_i,$$

and $$L_i[f] = f_i;$$

hence $$L_i[T_n] = L_i[f], \qquad (i = 0, 1, \cdots, n);$$

i. e., equations (2) are satisfied, so that *the minimizing set* $\{T_n(x)\}$ *is identical with the best approximating set* $\{s_n(x)\}$.

As a final example, consider the following class of methods: Let

$$J(t) \sim a_1 t + a_2 t^2 + \cdots, \qquad (a_1 \neq 0)$$

be a formal power series, generating the operator

$$J[u(x)] = a_1 u'(x) + a_2 u''(x) + \cdots.$$

Now define

$$M_0[u(x)] \equiv u(x),\ M_1[u(x)] \equiv J[u(x)], \cdots, M_n[u(x)] \equiv J^n[u(x)], \cdots,$$

giving the method $\mathfrak{M}$:

$$L_n[u(x)] = \{M_n[u(x)]\}_{x=0}.$$

We have already pointed out the cases $J(t) = t$, giving $J[u] = du/dx$, and $J(t) = e^t - 1$, giving $J[u] = u(x+1) - u(x)$. For the general J there will exist a set of best approximating polynomials $\{\Phi_n(x)\}$, which is of considerable interest in the study of functional equations based on the operator $J[u(x)]$. This aspect of these polynomial sets will not concern us here, but we wish to point out an interesting recurrence relation among the polynomials of the set $\{\Phi_n(x)\}$. It is this: [10]

$$J[\Phi_n(x)] = \Phi_{n-1}(x), \qquad (n = 1, 2, \cdots).$$

Let us turn once again to the Widder case. We can write

[10] The following two particular cases are well-known:

$$\frac{d}{dx}\left[\frac{x^n}{n!}\right] = \frac{x^{n-1}}{(n-1)!};$$

$$\Delta\left[\frac{x(x-1)\cdots(x-n+1)}{n!}\right] = \frac{x(x-1)\cdots(x-n+2)}{(n-1)!}.$$

$$\Omega_n(x) = \frac{x^n}{n!}[1 + h_n(x)] = \frac{x^n}{n!} + c_{n,n+1}\frac{x^{n+1}}{(n+1)!} + c_{n,n+2}\frac{x^{n+2}}{(n+2)!} + \cdots;$$

or, on setting $\Phi_n(x) = x^n/n!$:

$$\Omega_n(x) = \Phi_n(x) + c_{n,n+1}\Phi_{n+1}(x) + c_{n,n+2}\Phi_{n+2}(x) + \cdots.$$

That is, $\{\Phi_n\}$ is a basic set for $\{\phi_n(x) = x^n\}$, and $\{\Omega_n\}$ is a basic set for some sequence, say $\{\omega_n(x)\}$; and $\Omega_n(x)$ has the above expression in terms of the set $\{\Phi_n\}$. This is a fairly general phenomenon, as the following theorem will show.

Let $\mathfrak{M}$ be a given method, and $\{\phi_n(x)\}$ a sequence relative to which $\mathfrak{M}$ is non-singular. Then, as we have seen, there exists a unique basic set $\{\Phi_n\}$.

THEOREM 2. *Let $\{\phi_n(x)\}$, $\{\omega_n(x)\}$ be two sequences for which $\mathfrak{M}$ is non-singular, and let $\{\Phi_n(x)\}$, $\{\Omega_n(x)\}$ be their basic sets. If*

(i) *each $\omega_n(x)$ has a convergent Φ_n-expansion in a region $\mathfrak{R}$:*

$$\omega_n(x) = \sum_{i=0}^{\infty} a_{ni}\Phi_i(x); \tag{12}$$

(ii) *the operators $L_0, L_1, \cdots$ (which define $\mathfrak{M}$) are term-wise applicable to the above expansions:*

$$L_m[\omega_n(x)] = \sum_{i=0}^{\infty} a_{ni}L_m[\Phi_i(x)]; \tag{13}$$

then the set $\{\Omega_n(x)\}$ has the form (convergent in $\mathfrak{R}$)

$$\Omega_n(x) = \Phi_n(x) + c_{n,n+1}\Phi_{n+1}(x) + c_{n,n+2}\Phi_{n+2}(x) + \cdots. \tag{14}$$

We observe first that since Ω_n is a linear combination of $\omega_0, \cdots, \omega_n$, it possesses a Φ_n-expansion convergent in $\mathfrak{R}$: $\Omega_n(x) = \sum_{i=0}^{\infty} c_{ni}\Phi_i(x)$. Again, condition (ii) allows term-wise operation by L_m on this series: $L_m[\Omega_n] = \sum_{i=0}^{\infty} c_{ni}L_m[\Phi_i]$. Now $\{\Omega_n\}$, $\{\Phi_n\}$ are basic sets, so (5) holds for them. Taking $m = 0, 1, \cdots, n$, this yields the relations

$$c_{n0}L_m[\Phi_0] + c_{n1}L_m[\Phi_1] + \cdots + c_{nm}L_m[\Phi_m] = \begin{cases} 0, & (m = 0, 1, \cdots, n-1), \\ 1, & (m = n); \end{cases}$$

which, on recalling that $L_m[\Phi_m] = 1$, gives the following values for c_{ni}: $c_{n0} = c_{n1} = \cdots = c_{n,n-1} = 0$, $c_{nn} = 1$. Hence (14) holds.

2. *Convergence in the Widder case.*[11] We are here concerned with the convergence properties of Ω_n-series, where

[11] We have already remarked that in this section we shall lighten one of Widder's conditions. We add that the method used here is more direct than that of Widder.

(15) $\Omega_n(x) = (x^n/n!)[1 + h_n(x)] = (x^n/n!)\Theta_n(x), \qquad h_n(0) = 0,$

with $h_n(x)$ analytic in $|x| \leqq R$.

THEOREM 3. *Suppose constants c, N, β_n exist such that*

$$\text{(i)} \qquad 0 < c \leqq |\Theta_n(x)| \leqq \beta_n,$$

uniformly in $|x| \leqq R$ for all $n > N$, with[12]

$$\text{(ii)} \qquad \limsup_{n=\infty} \beta_n{}^{1/n} \leqq 1.$$

If the series

$$(16) \qquad f(x) = \sum_{n=0}^{\infty} f_n \Omega_n(x)$$

converges for a single point $x = \xi$ in $|x| \leqq R$, it converges uniformly and absolutely in every closed region lying[13] *in $|x| < |\xi|$, thus representing an analytic function in $|x| < |\xi|$.*

For:

$$\sum_0^\infty |f_n\Omega_n(x)| = \sum_0^\infty |f_n\Omega_n(\xi)|\,|\Omega_n(x)/\Omega_n(\xi)| \leqq A(x) + (M/c)\sum_{n=N}^{\infty} \beta_n |x/\xi|^n,$$

where $A(x)$ is the sum of the absolute values of the first N terms, and M is a bound (which exists) of $|f_n\Omega_n(\xi)|$. $\Omega_n(\xi)$ can vanish only if $\xi = 0$ or $\Theta_n(\xi) = 0$. Now the theorem is vacuously true if $\xi = 0$; and $\Theta_n(\xi) \neq 0$ by virtue of (i). Hence we can assume that $\Omega_n(\xi) \neq 0$; and the indicated division is possible. Since[14] $\limsup \beta_n{}^{1/n} \leqq 1$, the last series converges uniformly and absolutely in every closed region in $|x| < |\xi|$, and this is then true of the original series.

It is seen that condition (i), although applying *throughout* $|x| \leqq R$, is used only at the point $x = \xi$. Now it may happen that for some points in $|x| \leqq R$ a number c (depending on the point) exists, and for other points it does not. This suggests strengthening Theorem 3 as follows:

THEOREM 3′. *Let Σ be the set of those points $x = \xi$ in $|x| \leqq R$ for which c, N, β_n exist (as functions of ξ) such that*

A class of Ω_n-series (that arose in the study of some linear differential equations) is the $\mathfrak{D}_n$-series of Transactions, *American Mathematical Society*, vol. 35 (1933), pp. 184-214.

[12] Incidentally, (ii) combined with $c \leqq \beta_n$ gives $\limsup \beta_n{}^{1/n} = 1$.

[13] It follows that if the region of convergence does not go outside of $|x| \leqq R$, then it is a circular region.

[14] If condition (ii) is replaced by (ii′) $\limsup \beta_n{}^{1/n} \leqq K$, then the region for which convergence can be asserted is $|x| < |\xi|/K$.

(i) $$0 < c(\xi) \leqq |\Theta_n(\xi)| \leqq \beta_n(\xi),$$
with
(ii) $$\limsup_{n=\infty} [\beta_n(\xi)]^{1/n} \leqq 1.$$

If series (16) *converges for a single point* $x = \xi$ *in* Σ, *it converges uniformly and absolutely in every closed region lying in* $|x| < |\xi|$, *to an analytic function.*

The proof of Theorem 3 applies to 3′.

LEMMA 2. *A function cannot have two* $\Omega_n(x)$*-expansions uniformly convergent in an open region* $\mathfrak{R}$ *containing the point* $x = 0$.

For on subtracting we should have $0 = \Sigma a_n\Omega_n(x)$, uniformly convergent in $\mathfrak{R}$. By successsive term-wise differentiations (which are permissible) at $x = 0$, we find that $(a_n = 0,\ n = 0, 1, \cdots)$.

Let us try to develop the function $1/(t - x)$ (t a parameter) in an Ω_n-series. Assume that

(17) $$1/(t - x) = \sum_{n=0}^{\infty} L_n(t)\Omega_n(x).$$

By (formal) term-wise differentiation and setting $x = 0$, we get

(18) $$\begin{aligned} 0!/t &= L_0(t) \\ 1!/t^2 &= L_0(t)h'_0(0) + L_1(t) \\ &\cdots\cdots\cdots\cdots \\ n!/t^{n+1} &= L_0(t)h_0^{(n)}(0) + L_1(t)\binom{n}{1}h_1^{(n-1)}(0) \\ &\quad + \cdots + L_{n-1}(t)\binom{n}{n-1}h'_{n-1}(0) + L_n(t) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

thus determining the functions $\{L_n(t)\}$. $L_n(t)$ is, in fact, a polynomial in $1/t$ of degree $n + 1$.

Define λ_n as the maximum of $|h_n(x)|$ in $|x| \leqq R$:

(19) $$|h_n(x)| \leqq \lambda_n, \qquad |x| \leqq R.$$
Then
(20) $$|h_n^{(m)}(0)| \leqq m!\lambda_n/R^m.$$

Let r be any positive number, and let $\rho = \min(r, R)$. A simple application of (20) to (18) yields the inequalities

$$|L_0(t)| \leqq 0!/\rho; \qquad |L_1(t)| \leqq (1!/\rho^2)(1 + \lambda_0);$$
$$|L_2(t)| \leqq (2!/\rho^3)(1 + \lambda_0)(1 + \lambda_1),$$

uniform in $|t| \geqq r$; and a straightforward induction proof gives

LEMMA 3. *For all n, and uniformly in $|t| \geqq r$, where r is any positive number and $\rho = min(r, R)$,*

(21) $$|L_n(t)| \leqq (n!/\rho^{n+1})(1+\lambda_0)(1+\lambda_1)\cdots(1+\lambda_{n-1}).$$

Then, for $|t| \geqq r$, $|x| \leqq R$,

$$\sum_0^\infty |L_n(t)\Omega_n(x)| = \sum_0^\infty |L_n(t)\frac{x^n}{n!}[1+h_n(x)]| << \frac{1}{\rho}\sum_0^\infty \{\prod_{i=0}^n (1+\lambda_i)\}\left|\frac{x}{\rho}\right|^n.$$

If u_n is the n-th term of the series on the right, $u_{n+1}/u_n = (1+\lambda_{n+1})|x/\rho|$.

THEOREM 4. *Consider the series $\sum_0^\infty L_n(t)\Omega_n(x)$, where $\{L_n(t)\}$ is given by* (18), *and $|h_n(x)| \leqq \lambda_n$, $|x| \leqq R$. If lim sup $\lambda_n = K$ (finite), the series converges uniformly and absolutely in $|x| \leqq l$, $|t| \geqq r$, where r is any positive number, $\rho = min(r, R)$, and l is any positive number less than $\rho/(K+1)$; and the series represents, in this region, the function $1/(t-x)$.*

The convergence properties stated follow from the preceding relations. Let $H(x,t)$ be the sum of the series; it is analytic in x and t in $|x| \leqq l$, $|t| \geqq r$. Term-wise differentiation in x (which is permissible) with $x = 0$ gives

$$\left\{\frac{\partial^n H(x,t)}{\partial x^n}\right\}_{x=0} = n!/t^{n+1}, = \left\{\frac{\partial^n[1/(t-x)]}{\partial x^n}\right\}_{x=0}, \quad (n = 0, 1, \cdots);$$

hence H coincides with $1/(t-x)$.

Especially interesting is the case $K = 0$, in which case lim sup λ_n $= \lim \lambda_n = 0$:

THEOREM 5. *If $\lim \lambda_n = 0$, then series* (17) *is valid, converging uniformly and absolutely in $|x| \leqq l$, $|t| \geqq r$, where $r > 0$ is arbitrary, $\rho = min(r, R)$, and l is any positive number less than ρ.*

THEOREM 6. *If*[15] *$\lim \lambda_n = 0$, every function $f(x)$, analytic about $x = 0$, possesses an Ω_n-expansion. If the distance from $x = 0$ to the nearest singularity of $f(x)$ is a, and if $\sigma = min(a, R)$, this expansion is uniformly and absolutely convergent in $|x| \leqq \tau$ where τ is any positive number $< \sigma$, and the coefficients of the expansion are given by*

(22) $$f(x) = \sum_0^\infty f_n\Omega_n(x), \qquad f_n = (1/2\pi i)\int_C f(t)L_n(t)\,dt,$$

[15] If lim sup $\lambda_n = K$ (finite), every $f(x)$, analytic about $x = 0$, has an Ω_n-expansion, but with reduced radius of convergence.

C being any contour around $t=0$ and within $|t|<\sigma$. Moreover, $f(x)$ has only one Ω_n-expansion.

The convergence property follows at once on multiplying (17) through by $f(t)$ and integrating around C, using the Cauchy integral formula. All that remains is the uniqueness proof. From $\lim \lambda_n = 0$ follows the existence of β_n satisfying the conditions of Theorem 3. Hence if $f(x)$ possesses an expansion different from (22) and convergent in at least one point $x=\xi\neq 0$, it converges uniformly in an open region containing $x=0$. Lemma 2 now applies, to give us a contradiction; hence there is uniqueness.

COROLLARY. *The sets $\{\Omega_n\}$, $\{L_n\}$ are normal-orthogonal on C, any contour in $|t|<R$, surrounding $t=0$:*

$$(1/2\pi i)\int_C \Omega_m(t)L_n(t)\,dt = \begin{cases} 0, & (m\neq n), \\ 1, & (m=n). \end{cases}$$

For, $\Omega_m(x)$ is analytic in $|x|\leqq R$ and therefore possesses a *unique* Ω_n-expansion, the coefficient of $\Omega_n(x)$ being the above integral. Hence normality and orthogonality hold.

Theorem 6 can be given a different form:

THEOREM 7. *Let the condition $\lim \lambda_n = 0$ be replaced by the following condition: $\Theta_n(x) = 1 + h_n(x)$ converges uniformly in $|x|\leqq R$ to a function $M(x)$ which is nowhere zero in $|x|\leqq R$. Then the conclusion of Theorem 6 is valid with the modification that f_n is now given by*

$$f_n = (1/2\pi i)\int_C [f(t)/M(t)]L^*_n(t)\,dt, \tag{23}$$

where $\{L^*_n(t)\}$ is defined by (18) with $h_n(x)$ replaced by $g_n(x)$, the latter defined by

$$1 + h_n(x) = M(x)[1 + g_n(x)]. \tag{24}$$

Clearly, $g_n(x)\to 0$ uniformly, $|x|\leqq R$. The series $f(x) = \sum_0^\infty f_n\Omega_n(x)$ is identical with the series $f(x)/M(x) = \sum_0^\infty f_n(x^n/n!)[1+g_n(x)]$, and since $\{\lambda_n\}$ exists such that $|g_n(x)|\leqq\lambda_n$, $\lim \lambda_n = 0$, the second series expansion is valid by Theorem 6. The theorem now follows from the fact that the class of functions $\{f(x)\}$ analytic about $x=0$ coincides with the class of functions $\{f(x)/M(x)\}$.

3. *Extension of the Birkhoff theory.* The method of this section is

adapted from the work of Birkhoff, as was mentioned.[16] We start with a set of functions $\{u_n(x)\}$, with given convergence properties, and seek to determine the convergence properties of a second set of functions $\{v_n(x)\}$, related to $\{u_n(x)\}$ only quantitatively. We shall introduce certain assumptions labelled Condition A, B, C; and it is to be understood that once a Condition has been stated, it is to hold from then on to the end of the section.

Consider a sequence of functions $\{u_n(x)\}$ satisfying Condition A:

(i) *$u_n(x)$ is analytic in the interior $\mathfrak{S}$ of a rectifiable, simple closed curve C, and is continuous in $\mathfrak{S} + C$.*

(ii) *The function $1/(t-x)$, t a parameter, possesses a u_n-expansion*

$$1/(t-x) = \sum_{n=0}^{\infty} L_n(t)u_n(x), \tag{24}$$

which is uniformly convergent in x and t for t on C and x on any closed point set in $\mathfrak{S}$; and the functions $\{L_n(t)\}$ are continuous on C.[17]

Corollary. *Every function $f(x)$ that is analytic in $\mathfrak{S}$ and continuous in $\mathfrak{S} + C$, has a $u_n(x)$-expansion, uniformly convergent on every closed point set in $\mathfrak{S}$:*

$$(25)\quad f(x) = \sum_{n=0}^{\infty} f_n u_n(x), \qquad (26)\quad f_n = (1/2\pi i)\int_C f(t)L_n(t)\,dt.$$

Corollary. *If $\Sigma \mid L_n(t)u_n(x) \mid$ converges uniformly, t on C and x on any closed set in $\mathfrak{S}$, then series (25) converges absolutely, x in $\mathfrak{S}$.*

Now consider the set $\{v_n(x)\}$, which is to be "close" to the set $\{u_n(x)\}$ in a sense to be defined. *We assume that $v_n(x)$ is analytic in $\mathfrak{S}$ and continuous in $\mathfrak{S} + C$.* Suppose we have the expansion

$$f(x) = \sum_{0}^{\infty} \phi_n v_n(x). \tag{27}$$

---

[16] A number of papers have been written on subjects related to the work of Widder and of Birkhoff. References are to be found in Widder's paper and also in: Walsh, *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 53-57. In this section we do not emphasize generality of statement in our theorems. Rather, we aim to secure a comprehensive body of theorems that are symmetric (i. e., interchangeable) in the two sets of functions $\{u_n\}$, $\{v_n\}$; and that can be utilized in treating convergence of series of best approximating "polynomials."

[17] In Birkhoff's case, $u_n(x) = x^n$, so that Condition A holds when C is any circle with center at $x = 0$.

In analogy with (25, 26), we are led to consider the possibility [18] of defining the coefficients ϕ_n by

$$\phi_n = (1/2\pi i) \int_C g(t) L_n(t)\, dt, \tag{28}$$

$g(t)$ to be determined.

We see from (28) and (25, 26) that

$$g(x) = \sum_{n=0}^{\infty} \phi_n u_n(x). \tag{29}$$

Substitution of (28) into (27) yields

$$f(x) = (1/2\pi i) \int_C \{ \sum_{n=0}^{\infty} v_n(x) L_n(t) \} g(t)\, dt. \tag{30}$$

This integral equation is not well-adapted to determine $g(t)$. We can obtain an equation of the second kind by subtracting from (30) the relation

$$g(x) = \frac{1}{2\pi i} \int_C \frac{g(t)}{t-x}\, dt = \frac{1}{2\pi i} \int_C \{ \sum_{n=0}^{\infty} u_n(x) L_n(t) \} g(t)\, dt. \tag{31}$$

In fact, we then have

$$f(x) = g(x) + (1/2\pi i) \int_C K(x, t) g(t)\, dt \tag{32}$$

where

$$K(x, t) = \sum_{n=0}^{\infty} [v_n(x) - u_n(x)] L_n(t). \tag{33}$$

We now assume

Condition B. *Series* (33) *converges uniformly for x in $\mathfrak{I} + C$ and t on C, and*

$$| K(x, t) | < 2\pi/l$$

for x and t on C, where $l =$ length of C.

It is Condition B that is the test of $\{v_n(x)\}$ being "*close*" to $\{u_n(x)\}$.

COROLLARY. *$K(x, t)$ is analytic in x in the region $\mathfrak{I}$ for each t on C, and is continuous in x and t for t on C and x in $\mathfrak{I} + C$.*

We observe that the formal work from (27) to (33) is valid *if we work backwards*; i. e., *given* $g(x)$, assumed to be analytic in $\mathfrak{I}$ and continuous in $\mathfrak{I} + C$; then (28), (29), (31) hold, and (29) is uniformly convergent for x on any closed set in $\mathfrak{I}$. If now $f(x)$ *is defined* by (32), then $f(x)$ is seen

[18] Our argument is purely formal until our conclusions are stated and proved.

to be analytic in $\mathfrak{D}$ and continuous in $\mathfrak{D}+C$; and by combining (31) and (32), then (30) holds, the series within the brace being uniformly convergent in x and t for t on C and x on any closed set in $\mathfrak{D}$. (30) may then be integrated term-wise, yielding series (27), which is uniformly convergent for x on any closed set in $\mathfrak{D}$.

Our aim, however, is to start with $f(x)$ and determine $g(x)$. In (32) let x be chosen on C. As x and t traverse C they are functions of the arc length (measured from some point on C):

$$x=\Theta(s), \qquad t=\Theta(\sigma).$$

Our hypothesis on C assures us of the existence of $d\Theta(\sigma)/d\sigma$ almost everywhere.[19] Point x being on C, let us set

$$F(s)=f(\Theta(s)),\ G(s)=g(\Theta(s)),\ K^*(s,\sigma)=(1/2\pi i)K(\Theta(s),\Theta(\sigma))\Theta'(\sigma).$$

Equation (32) then reduces to the equivalent form

$$F(s)=G(s)+\int_0^l K^*(s,\sigma)G(\sigma)\,d\sigma. \tag{32$'$}$$

$F(s)$ is continuous; and so is $K^*(s,\sigma)$ except on a set of measure zero (due to the possible non-existence of $\Theta'(\sigma)$), where it can be defined so as to be bounded for s, σ in $0\leqq s,\ \sigma\leqq l$. The Fredholm theory can be applied. Because of the second part[20] of Condition B, the Neumann series for a solution of (32′) converges uniformly, so that $\lambda=1$ is not a characteristic number. Hence (32′) has a unique solution $G(s)$; and it is continuous, $0\leqq s\leqq l$.

This continuous function $G(s)$ defines a continuous function $g(x)$ (x on C) satisfying (32); and $g(x)$ is a unique solution of (32). We now extend the definition of $g(x)$ to $\mathfrak{D}$ by means of (32), where x is now in $\mathfrak{D}$. $g(x)$ is seen to be analytic in $\mathfrak{D}$. Suppose x, in $\mathfrak{D}$, approaches a point α of C. The functions $f(x)$, $(1/2\pi i)\int_C K(x,t)g(t)\,dt$ being continuous in $\mathfrak{D}+C$, it follows from (32) that $g(x)\to g(\alpha)$, where $g(\alpha)$ is the value, at $x=\alpha$, of the unique solution of (32) *for x on C*. This gives us

[19] And at such points where $\Theta'(\sigma)$ fails to exist, the difference quotient is nevertheless bounded: $|\Delta\Theta/\Delta\sigma|\leqq 1$. Where Θ' does exist, it has the value $\Theta'(\sigma)=e^{i\theta}$ where θ is the angle which the tangent to C (at the point σ) makes with the real axis.

[20] Cf. Whittaker and Watson, *Modern Analysis*, 4th ed., pp. 221-222. A remark of Birkhoff (*loc. cit.*) is apropos here: It is not necessary that $|K(x,t)|$ be less than $2\pi/l$. *All* the work of this section will hold if we merely assume that in the integral equations (32) and (40), $\lambda=1$ is not a characteristic number.

LEMMA 4. *To every function $f(x)$, analytic in $\mathfrak{I}$ and continuous in $\mathfrak{I} + C$, there corresponds a unique solution $g(x)$ of (32); and $g(x)$ is also analytic in $\mathfrak{I}$ and continuous in $\mathfrak{I} + C$.*

Having obtained the function $g(x)$, the observation made after Condition B, on going from $g(x)$ to $f(x)$, enables us to state

THEOREM 8. *Every function $f(x)$, analytic in $\mathfrak{I}$ and continuous in $\mathfrak{I} + C$, has a $v_n(x)$-expansion*

$$(27)\quad f(x) = \sum_{n=0}^{\infty} \phi_n v_n(x), \qquad (28)\quad \phi_n = (1/2\pi i) \int_C g(t) L_n(t)\, dt,$$

uniformly convergent on every closed set in $\mathfrak{I}$. In (28), $g(t)$ is the function of Lemma 4.

COROLLARY. *If series $\Sigma \mid u_n(x) L_n(t) \mid$ converges uniformly for x in $\mathfrak{I}$ and t on C, and series $\Sigma \mid [v_n(x) - u_n(x)] L_n(t) \mid$ converges uniformly for x in $\mathfrak{I} + C$ and t on C, then series (27) (with ϕ_n given by (28)) converges absolutely for all x in $\mathfrak{I}$.*

For: We have $f(x) = \Sigma \phi_n v_n(x)$, $g(x) = \Sigma \phi_n u_n(x)$, so that

$$f - g = \Sigma (1/2\pi i) \int_C g(t) L_n(t) [v_n(x) - u_n(x)]\, dt.$$

By hypothesis this series converges absolutely. Again, from the second Corollary to Condition A, $g = \Sigma (1/2\pi i) \int_C g(t) L_n(t) u_n(x)\, dt$ converges absolutely. Hence on adding, (27) converges absolutely, x in $\mathfrak{I}$.

The functions $\{L_n(t)\}$ are defined only on C, where they are continuous. We can extend their definition *to $\mathcal{E}$, the region exterior to C:*

DEFINITION. *For z in $\mathcal{E}$, $L_n(z)$ is defined to be*

$$(34)\qquad L_n(z) = \frac{-1}{2\pi i} \int_C \frac{L_n(t)}{t - z}\, dt.$$

COROLLARY. *$L_n(z)$ is analytic*[21] *in $\mathcal{E}$, and $L_n(\infty) = 0$.*

THEOREM 9. *The series*

$$(35)\qquad 1/(z - x) = \sum_{n=0}^{\infty} L_n(z) u_n(x),$$

[21] There is no reason for supposing, without further hypotheses, that L_n, defined in $\mathcal{E} + C$ by (34) and by Condition A, is continuous in $\mathcal{E} + C$. Later, when we *do* have a further condition, this assertion can be made. (See Theorem 20.)

holds uniformly in x and z for x on any closed set in $\mathfrak{I}$ and z on any "closed" set[22] *in $\mathcal{E}$.*

To show this, observe that for z in $\mathcal{E}$, $1/(z-x)$ is analytic (in x) in $\mathfrak{I}$ and continuous in $\mathfrak{I}+C$, so that (Corollary to Condition A) it has a uniformly convergent $u_n(x)$-expansion:

$$\frac{1}{z-x}=\sum_0^\infty \phi_n(z)u_n(x), \qquad \phi_n(z)=\frac{1}{2\pi i}\int_C\left(\frac{1}{z-t}\right)L_n(t)\,dt.$$

From (34) we see that $\phi_n(z)=L_n(z)$. Since $\sum\limits_0^\infty u_n(x)L_n(t)$ converges uniformly, x on a closed set in $\mathfrak{I}$ and t on C, therefore term-wise integration (after multiplication by $1/(z-t)$) is permissible, the resulting series being uniformly convergent for x and z in the regions stated.

THEOREM 10. *If $f(z)$ is analytic in $\mathcal{E}+C$, it has the $L_n(z)$-expansion*

$$(36)\quad f(z)-f(\infty)=\sum_{n=0}^\infty \alpha_n L_n(z), \qquad (37)\quad \alpha_n=\frac{1}{2\pi i}\int_C u_n(t)f(t)\,dt,$$

uniformly convergent for z on any "closed" set in $\mathcal{E}$.

For: We can find a rectifiable simple closed curve J inside C such that $f(z)$ is analytic on J and exterior to J. Then, by (35),

$$f(z)-f(\infty)=\frac{-1}{2\pi i}\int_J\frac{f(x)}{x-z}\,dx=\sum_0^\infty\left\{\frac{1}{2\pi i}\int_J u_n(x)f(x)\,dx\right\}L_n(z),$$

the series being uniformly convergent for z on a "closed" set in $\mathcal{E}$. Now J can be chosen as close to C as we wish; whence it follows, from the continuity of $u_n(x)$ and $f(x)$ in the closed region consisting of J, C and the ring that they bound, that in the coefficient of $L_n(z)$ the curve of integration J may be replaced by C without altering values. That is, (37) holds.

Consider again equation (33). The series being uniformly convergent for x in $\mathfrak{I}+C$ and t on C, we may multiply through by $1/(t-z)$, z in $\mathcal{E}$, and integrate term-wise:

$$\begin{aligned}\frac{-1}{2\pi i}\int_C\frac{K(x,t)}{t-z}\,dt&=\frac{-1}{2\pi i}\int_C\sum_0^\infty[v_n(x)-u_n(x)]\frac{L_n(t)}{t-z}\,dt\\&=\sum_0^\infty[v_n(x)-u_n(x)]\left(\frac{-1}{2\pi i}\int_C\frac{L_n(t)}{t-z}\,dt\right),\end{aligned}$$

[22] By a "closed" set in $\mathcal{E}$ we shall mean both the usual closed set and also any unbounded set in $\mathcal{E}$ (including $z=\infty$), the important feature being that the set is at a positive distance from C.

the resulting series being uniformly convergent for x in $\mathfrak{J} + C$, and z on any "closed" set in $\mathcal{E}$. Now by (34), the parenthesis has the value $L_n(z)$. This enables us to extend the definition of K to $\mathcal{E}$:

LEMMA 5. *The series*

$$K(x,z) = \sum_0^\infty [v_n(x) - u_n(x)]L_n(z) \tag{37}$$

converges uniformly in x and z for x in $\mathfrak{J} + C$ and z on any "closed" set in $\mathcal{E}$, so that $K(x,z)$ is analytic in x and z for x in $\mathfrak{J}$ and z in $\mathcal{E}$. Moreover,

$$K(x,z) = \frac{-1}{2\pi i}\int_C \frac{K(x,t)}{t-z}\,dt, \tag{38}$$

where $K(x,t)$ is given by series (33); *and $K(x,\infty) = 0$.*

In the expansion (27), the coefficients ϕ_n are given in terms of $g(x)$. The question arises if we can express ϕ_n directly in terms of $f(x)$:

$$\phi_n = (1/2\pi i)\int_C f(t)M_n(t)\,dt, \tag{39}$$

where the functions $M_n(t)$ are to be determined. Since (32) holds for x in $\mathfrak{J} + C$, we may substitute for $f(t)$ in (39) its value as given by (32). On further using the relation $\phi_n = (1/2\pi i)\int_C g(t)L_n(t)\,dt$, this gives

$$\text{(a)}\quad (1/2\pi i)\int_C g(t)\{L_n(t) - M_n(t) - (1/2\pi i)\int_C K(w,t)M_n(w)\,dw\}\,dt = 0.$$

Now (a) is to hold for all g, and we want M_n to be independent of f (and therefore of g). This suggests that we set the brace equal to zero:

$$L_n(t) = M_n(t) + (1/2\pi i)\int_C K(w,t)M_n(w)\,dw. \tag{40}$$

This integral equation can be thrown into "real" form, as was (32). The resulting kernel is $K^{**}(s,\sigma) = (1/2\pi i)K(\Theta(\sigma),\Theta(s))\Theta'(\sigma)$, so, as was the case with (32), (40) has a unique solution $M_n(t)$ (t on C), and $M_n(t)$ is continuous.

THEOREM 11. *In Theorem* 8, *the coefficients ϕ_n can also be expressed by* (39), *where $M_n(t)$ is the unique and continuous solution of* (40), *t on C.*

To see this, we observe first that (a) is satisfied. On using (28) and (32), (a) simplifies to

$$0 = \phi_n - (1/2\pi i)\int_C M_n(t)[g(t) + (1/2\pi i)\int_C K(t,w)g(w)\,dw]\,dt$$
$$= \phi_n - (1/2\pi i)\int_C M_n(t)f(t)\,dt,$$

which is (39).

By means of (40), with t replaced by z, we can extend the definition of M_n:

DEFINITION. *For z in $\mathcal{E}$,*

(41) $$M_n(z) = L_n(z) - (1/2\pi i)\int_C K(t,z)M_n(t)\,dt,$$

$M_n(t)$ being the unique and continuous solution of (40).

COROLLARY. *$M_n(z)$ is analytic in z, with $M_n(\infty) = 0$; and for z in $\mathcal{E}$,*

(42) $$M_n(z) = \frac{-1}{2\pi i}\int_C \frac{M_n(t)}{t - z}\,dt.$$

We need only establish (42). If we multiply (40) through by $[(-1)/2\pi i] \times [1/(t-z)]$ and integrate over C, and use (34) and (38), we get

$$L_n(z) = \frac{-1}{2\pi i}\int_C \frac{M_n(t)}{t - z}\,dt + \frac{1}{2\pi i}\int_C K(w,z)M_n(w)\,dw.$$

Comparison with (41) then yields (42).

THEOREM 12. *The expansion*

(43) $$1/(z - x) = \sum_0^\infty v_n(x)M_n(z)$$

is uniformly convergent in x and z for x on any closed set in $\mathfrak{I}$ and z on any "closed" set in $\mathcal{E}$.

For: In (32), choose $f(x) = f(x,z) = 1/(z - x)$, z in $\mathcal{E}$. Then $g(x) = g(x,z)$ is defined by (32) to be analytic in x and z for x in $\mathfrak{I}$ and z in $\mathcal{E}$; and is continuous for x in $\mathfrak{I} + C$ and z in $\mathcal{E}$. If we multiply series (24) through by $(1/2\pi i)\cdot[1/(z - t)]$ and integrate term-wise around C, we observe that $g(x,z)$ has a $u_n(x)$-expansion (cf. (29)) that is uniformly convergent in x and z, x on any closed set in $\mathfrak{I}$ and z on any "closed" set in $\mathcal{E}$:

$$g(x,z) = \sum_{n=0}^\infty u_n(x)\phi_n(z), \qquad \phi_n(z) = (1/2\pi i)\int_C L_n(t)g(t,z)\,dt.$$

If we now combine (31, 32, 33), where $g(x) = g(x,z)$ (so that $f(x) = f(x,z) = 1/(z - x)$), and recall that (31) *converges uniformly in*

both x and z, then (30) is seen to hold, also uniformly convergent in both x and z. Term-wise integration gives (27), again uniformly convergent in x and z: $1/(z-x)=\sum_0^\infty \phi_n(z)v_n(x)$. It remains to identify the coefficient of $v_n(x)$. This coefficient is the coefficient ϕ_n given by (27) and (28); and by Theorem 11 this coefficient has the value given by (39):

$$\phi_n(z)=\frac{1}{2\pi i}\int_C \frac{M_n(t)}{z-t}\,dt.$$

Comparison with (42) shows that $\phi_n(z)=M_n(z)$, and the theorem is established.

From this follows (cf. Theorem 10)

THEOREM 13. *If $f(z)$ is analytic in $\mathcal{E}+C$, it has the $M_n(z)$-expansion*

$$(44)\ f(z)-f(\infty)=\sum_{n=0}^\infty \beta_n M_n(z)\,dz; \qquad (45)\ \beta_n=(1/2\pi i)\int_C v_n(t)f(t)\,dt,$$

uniformly convergent for z on any "closed" set in $\mathcal{E}$.

There is apparent, by this time, a duality between the sets $\{u_n\}$, $\{L_n\}$ on the one hand and the sets $\{v_n\}$, $\{M_n\}$ on the other. We shall now examine to what extent their rôles can be interchanged. If $H(x,t)$ exists, having the relation to $\{v_n\}$ that $K(x,t)$ has to $\{u_n\}$, we should expect it to be given by the series

$$(45)\qquad H(x,t)=\sum_{n=0}^\infty [u_n(x)-v_n(x)]M_n(t).$$

For the moment we shall put aside the problem of convergence of this series. If we substitute (40) into (33), we get the relation [23]

$$(46)\qquad H(x,t)+K(x,t)+(1/2\pi i)\int_C H(x,w)K(w,t)\,dw=0.$$

This is an integral equation for $H(x,t)$, with the same kernell $K(w,t)$ as in (40). Knowing the properties of $K(x,t)$, we can therefore state the following properties for $H(x,t)$:

LEMMA 6. *The function $H(x,t)$ defined by (46) is the only solution; it is continuous in x and t for x in $\mathcal{I}+C$ and t on C, and is analytic (in x) for x in $\mathcal{I}$ and for each t on C; and is continuous in x and t for x in $\mathcal{I}+C$ and t in $\mathcal{E}$, and is analytic (in x and t) for x in $\mathcal{I}$ and t in $\mathcal{E}$.*

---

[23] This is the well-known equation for kernel and resolvent kernel in integral equation theory.

THEOREM 14. *$K(x, t)$ is a resolvent kernel for the equation*

$$g(x) = f(x) + (1/2\pi i) \int_C H(x, t) f(t) \, dt. \tag{47}$$

For, in the right-hand side of (47), let f have the value given by (32). On simplifying we obtain $g(x) + (1/2\pi i) \int_C g(t) A(x, t) \, dt$, where $A(x, t)$ is the left-hand member of (46); i. e., $A(x, t) \equiv 0$, so that the right-hand side of (47) *does* equal $g(x)$, and (47) is satisfied.

Since equation (47) has the solution (32) for *every* function $g(x)$ that is continuous on C, it follows from the Fredholm theory that (47) always has a *unique* solution. Hence

COROLLARY 1. *For every $g(x)$, analytic in $\mathfrak{I}$ and continuous in $\mathfrak{I} + C$, equation* (47) *has a unique solution $f(x)$. This solution is analytic in $\mathfrak{I}$ and continuous in $\mathfrak{I} + C$.*

It is also an immediate consequence of the uniqueness of solutions of both (32) and (47) that

COROLLARY 2. *$H(x, t)$ is a resolvent kernel of* (32); *i. e., that the unique solution of* (32) *is furnished by* (47).

If in (47) we set $f(x) \equiv K(x, w)$, we find on using (46) that $g(x) \equiv -H(x, w)$. Substituting these values of f and g into (32) then gives us the equation which is the twin of (46):

COROLLARY 3. *The functions H and K satisfy the equation*

$$H(x, t) + K(x, t) + (1/2\pi i) \int_C K(x, w) H(w, t) \, dw = 0, \tag{46'}$$

valid for x in $\mathfrak{I} + C$ and t in $\mathcal{E} + C$.

LEMMA 7. *The unique solution $M_n(t)$ of equation* (40) *is given by*

$$M_n(t) = L_n(t) + (1/2\pi i) \int_C H(w, t) L_n(w) \, dw. \tag{48}$$

This follows on using (46').

We can now establish the validity of (45).

THEOREM 15. *$H(x, t)$ has the expansion* (45), *which converges uniformly for x in $\mathfrak{I} + C$ and t on C.*

To show this, we first have series (33), uniformly convergent in the region stated. We therefore have (from (48)):

$$\begin{aligned}
&\sum_0^\infty [v_n(x) - u_n(x)]M_n(t) \\
&\quad = \sum_0^\infty [v_n(x) - u_n(x)]\{L_n(t) + (1/2\pi i)\int_C H(w,t)L_n(w)dw\} \\
&\quad = \sum_0^\infty [v_n(x) - u_n(x)]L_n(t) \\
&\qquad\qquad + (1/2\pi i)\int_C H(w,t)\{\sum_0^\infty [v_n(x) - u_n(x)]L_n(w)dw\},
\end{aligned}$$

the two series on the right converging uniformly in the region stated. Hence this same convergence property applies to the series on the left. There remains only to prove that this series has the value $-H(x,t)$. But this follows from (46′) since the right-hand side is $K(x,t) + (1/2\pi i)\int_C H(w,t)K(x,w)dw$.

Theorem 8 can be dualized:

THEOREM 16. *If $g(x)$ is analytic in $\mathfrak{J}$ and continuous in $\mathfrak{J} + C$, then we have*

(49) $g(x) = \sum_0^\infty \psi_n u_n(x)$, (50) $\psi_n = (1/2\pi i)\int_C f(t)M_n(t)dt$,

uniformly convergent on any closed in $\mathfrak{J}$. Here $f(x)$ is the solution of (47).

For: In (47) replace $H(x,t)$ by its uniformly convergent expansion (45). There results the equation

$$g(x) = f(x) + \Sigma[(1/2\pi i)\int_C f(t)M_n(t)dt][u_n(x) - v_n(x)],$$

the series being uniformly convergent on any closed set in $\mathfrak{J}$. But the series

$$f(x) = \Sigma[(1/2\pi i)\int_C f(t)M_n(t)dt]v_n(x)$$

has the same convergence property (Theorem 11); hence so has

$$\Sigma[(1/2\pi i)\int_C f(t)M_n(t)dt]u_n(x).$$

We thus have

$$g(x) = f(x) + \Sigma[(1/2\pi i)\int_C f(t)M_n(t)dt]u_n(x) - f(x),$$

from which (49) follows.

We have now an almost complete duality of $\{u_n\}$, $\{L_n\}$ and $\{v_n\}$, $\{M_n\}$. That it is not fully complete (at least so far as has been proved) is owing to

this lack; in Condition A we are not certain that (24) holds in the region stated, when u_n, L_n are replaced by v_n, M_n. What we do know, up to this point, is that (24) will hold (cf. (43)) *if t is in $\mathcal{E}$*, rather than on C; nor does Theorem 8 permit t to be on C. However, we can fill in the gap:

THEOREM 17. *The expansion*

$$1/(t-x) = \sum_{n=0}^{\infty} M_n(t) v_n(x) \tag{50}$$

is valid, and is uniformly convergent in x and t for x on any closed point set in $\mathfrak{J}$ and t on C.

To show this, we begin with the expansion

$$1/(t-x) = \sum_{0}^{\infty} L_n(t) u_n(x), \tag{a}$$

which has the convergence properties stated above. On multiplying through by $(1/2\pi i)H(t,w)$ we may integrate term-wise, the resulting series being likewise uniformly convergent:

$$\sum_{0}^{\infty} (1/2\pi i) \int_C H(w,t) \cdot u_n(x) L_n(w)\,dw. \tag{b}$$

Hence the series $\sum_{0}^{\infty} \{L_n(t) + (1/2\pi i)\int_C H(w,t) L_n(w)\,dw\} u_n(x)$ converges uniformly in x and t for x on any closed set in $\mathfrak{J}$ and t on C. But the brace equals $M_n(t)$ (cf. (48)); hence the series $\sum_{0}^{\infty} M_n(t) u_n(x)$ converges uniformly in x and t in the region stated. Now (b) simplifies to

$$(1/2\pi i) \int_C H(w,t) \cdot [1/(w-x)]\,dw = H(x,t),$$

so that

$$\sum_{0}^{\infty} M_n(t) u_n(x) = 1/(t-x) + H(x,t) = 1/(t-x) + \sum_{0}^{\infty} [u_n(x) - v_n(x)] M_n(t).$$

The two series are uniformly convergent in x and t for x on any closed set in $\mathfrak{J}$ and t on C. If then we subtract the first series from both members, we get (50), the series having the same convergence properties. This establishes the theorem.[24]

[24] There is another point concerning duality: In Condition B we have $|K(x,t)| < 2\pi/l$. Now we do not know that H satisfies the same condition. In fact, if we write $\max |K(x,t)| = 2\pi\sigma/l$, $\sigma < 1$, then from (46) all we know is that $\max |H(x,t)|$

We may sum up, in part, as follows: *If* $\{u_n\}$ *satisfies Conditions A and B, so*[25] *does* $\{v_n\}$.

The function $u_m(x)$, being analytic in $\mathfrak{A}$ and continuous in $\mathfrak{A}+C$, has a u_n-expansion (cf. (25)):

$$u_m(x)=\sum_{n=0}^{\infty} f_{mn}u_n(x), \qquad f_{mn}=(1/2\pi i)\int_C u_m(t)L_n(t)\,dt.$$

If a u_n-expansion is unique, then we have the biorthogonal property

$$(51) \qquad (1/2\pi i)\int_C u_m(t)L_n(t)\,dt=\begin{cases}0, & (m\neq n);\\ 1, & (m=n).\end{cases}$$

But Condition A does not insure uniqueness. For example, take $u_0(x)=1$; $u_n(x)=(x^{n-1}/n!)(x-n)$, $(n>0)$. It is readily shown[26] that if series $\Sigma c_n u_n(x)$ converges, the region of convergence is the interior of a circle, center the origin, reaching out to the nearest singularity of the function that is defined; and the convergence is uniform on any closed set within the circle of convergence. Moreover, every function, analytic about $x=0$, has a u_n-expansion; and the functions $\{L_n\}$ can be defined by

$$L_n(t)=-\sum_{i=0}^{n-1}(i!/t^{i+1}),\ (n>0); \qquad L_0(t)=-1/t.$$

Condition A is fulfilled on choosing C as any circle with center at $x=0$. But the function *zero* has the uniformly convergent expansion $0=\sum_0^{\infty} u_n(x)$, so there fails to be uniqueness.

If there is to be uniqueness, it must appear in our assumptions. We accordingly add the uniqueness

Condition C. *If zero has the expansion* $0=\sum_0^{\infty} a_n u_n(x)$, *uniformly convergent on every closed set in* $\mathfrak{A}$, *then* $a_n=0$, $(n=0,1,\cdots)$.

From this follows that a function $f(x)$ cannot have two distinct u_n-expansions, each uniformly convergent on every closed set in $\mathfrak{A}$. Consequently we have

---

$<[\sigma/(1-\sigma)](2\pi/l)$. But the only use made of the condition $|K|<2\pi/l$ is to secure uniqueness of the solutions of the integral equations with kernels $K(x,t)$ and $K(t,x)$. Hence we ought to establish this same uniqueness for the kernels $H(x,t)$ and $H(t,x)$. Now it is already known for $H(x,t)$ (cf. Theorem 14). That it also holds for $H(t,x)$ is easily shown.

[25] It being understood that in B the inequality $|K|<2\pi/l$ is replaced by the assumption of uniqueness of the solutions for the kernels $K(x,t)$ and $K(t,x)$.

[26] Compare the proof of Theorem 3.

LEMMA 8. *The biorthogonality relations* (51) *hold.*

LEMMA 9. *The sets* $\{v_n\}$, $\{M_n\}$ *are biorthogonal:*

$$(52) \qquad (1/2\pi i)\int_C v_m(t)M_n(t)\,dt = \begin{cases} 0, & (m \neq n); \\ 1, & (m = n). \end{cases}$$

Multiply equation (40) through by $u_m(t)$ and integrate over C:

$$\delta_{mn} = (1/2\pi i)\int_C M_n(t)u_m(t)\,dl + (1/2\pi i)^2\int_C\int_C K(w,t)M_n(w)u_m(t)\,dw\,dt.$$

If we replace $K(w,t)$ by its uniformly convergent expansion (33), this reduces to

$$\delta_{mn} = (1/2\pi i)\int_C M_n(t)u_m(t)\,dt + (1/2\pi i)\int_C [v_m(w) - u_m(w)]M_n(w)\,dw,$$

and on cancelling the first and third terms (whose sum is zero), there remains relation (52).

THEOREM 18. *Equations* (32) *and* (47) *are satisfied by taking* $f(x) = v_m(x)$, $g(x) = u_m(x)$.

To see this, let $g(x) = u_m(x)$ in (32) and replace $K(x,t)$ by its expansion (33). Then, using (51), $f(x) = u_m(x) + [v_m(x) - u_m(x)] = v_m(x)$.

We come now to $\{v_n\}$-uniqueness:

THEOREM 19. *If the v_n-expansion* $\sum_0^\infty c_n v_n(x)$ *converges uniformly in* $\mathfrak{D} + C$ (*the sum function* $f(x)$ *being therefore analytic in* $\mathfrak{D}$ *and continuous in* $\mathfrak{D} + C$), *then necessarily* $c_n = (1/2\pi i)\int_C f(t)M_n(t)\,dt$.

Multiply the series through by $M_s(x)$ and integrate over C:

$$(1/2\pi i)\int_C f(t)M_s(t)\,dt = \sum_{n=0}^{\infty} c_n \cdot (1/2\pi i)\int_C v_n(t)M_s(t)\,dt.$$

By biorthogonality, the series on the right reduces to c_s, thus establishing the theorem.

Suppose we make a temporary translation of the complex variable x so as to insure that the origin is within C. This will not affect any of the results already obtained. *Now* it will be true that

$$\frac{1}{2\pi i}\int_C \frac{dt}{t^m(t-x)} = 0, \qquad (m = 1, 2, \cdots);$$

whence from the uniformly convergent expansion (24), we obtain

$$0 = \sum_{n=0}^{\infty} \left\{ \frac{1}{2\pi i} \int_C \frac{L_n(t)}{t^m} \, dt \right\} u_n(x), \qquad (m = 1, 2, \cdots),$$

uniformly convergent for x on any closed set in $\mathfrak{I}$. By the uniqueness property, all the coefficients must vanish:

$$\text{(a)} \qquad \frac{1}{2\pi i} \int_C \frac{L_n(t)}{t^m} \, dt = 0, \qquad \begin{cases} (n = 0, 1, \cdots) \\ (m = 1, 2, \cdots). \end{cases}$$

Now for a given n, condition (a) is necessary and sufficient [27] that there exist a function (which will of necessity be $L_n(z)$), analytic in $\mathcal{E}$, and such that as $z \to t$ on C, $L_n(z) \to L_n(t)$.

It follows that $L_n(z)$ is continuous on $\mathcal{E} + C$, and analytic in $\mathcal{E}$. If we multiply (33) through by $(1/2\pi i) \cdot (1/t^m)$ and integrate around C, we see from (a) that

$$\frac{1}{2\pi i} \int_C \frac{K(x, t)}{t^m} \, dt = 0, \qquad (m = 1, 2, \cdots),$$

so that $K(x, z)$ also has the property that as z in $\mathcal{E}$ approaches a point t on C, then $K(x, z) \to K(x, t)$. That is, $K(x, z)$ is continuous for x in $\mathfrak{I} + C$ and z in $\mathcal{E} + C$. Finally, from (40) we get

$$\frac{1}{2\pi i} \int_C \frac{M_n(t)}{t^m} \, dt = 0, \qquad (m = 1, 2, \cdots),$$

so that $M_n(z)$ is continuous in $\mathcal{E} + C$; and therefore, also, $H(x, z)$. To sum up: [28]

THEOREM 20. *The functions $L_n(z)$, $M_n(z)$, $K(x, z)$, $H(x, z)$ are continuous in x and z for x in $\mathfrak{I} + C$ and z in $\mathcal{E} + C$; and their values for $z = t$ on C are respectively the known functions $L_n(t)$, $M_n(t)$, $K(x, t)$, $H(x, t)$.*

---

[27] If C is an analytic Jordan curve, this result holds if $\phi(t)$, the function given on the boundary (which in our case is $L_n(t)$) is continuous. (Walsh, *Transactions of the American Mathematical Society*, vol. 30 (1928), especially pp. 327 and 329.) If C is rectifiable, this same result holds if $\phi(t)$ is merely Lebesgue integrable (in which case the approach holds almost everywhere and must be non-tangential). (Priwaloff, *Comptes Rendus*, vol. 178 (1924), pp. 611-614.) In the Priwaloff case it is not clear (although probably true) that if $\phi(t)$ is continuous, then the approach holds everywhere on C. If this is not the case, we shall regard it as assumed that C satisfies the Walsh condition.

[28] If we now undo the translation that was made temporarily, none of these properties of continuity in $\mathcal{E} + C$ will be altered.

In some of our theorems we had to insist on z being on any "closed" set in $\mathcal{E}$, because we had not this last theorem. It will be clear that we can now amend some of the theorems, as follows:

COROLLARY. *In Theorem 9, uniform convergence maintains for z in $\mathcal{E}+C$; Theorems 12 and 17 combine to give uniform convergence for z in $\mathcal{E}+C$; and Lemma 5 holds uniformly for z in $\mathcal{E}+C$.*

If we were to consider interchanging the rôles of $\{u_n\}$ and $\{L_n\}$, or of $\{v_n\}$ and $\{M_n\}$, we would define two functions $A(z,x)$, $B(z,x)$ by the series

$$(53)\qquad \begin{aligned} A(z,x) &= \sum_{n=0}^{\infty}[M_n(z)-L_n(z)]u_n(x), \\ B(z,x) &= \sum_{n=0}^{\infty}[L_n(z)-M_n(z)]v_n(x). \end{aligned}$$

These functions are closely related to H and K. In fact we have the

COROLLARY. *The above series (53) converge uniformly in z and x for z on any "closed" set in $\mathcal{E}$ and x in $\mathfrak{S}+C$; and*

$$(54)\qquad A(z,x)=H(x,z); \qquad B(z,x)=K(x,z).$$

The convergence is immediate; and (54) follows from (33), (45) and (35), (43) (with reference to the preceding Corollary).

This is as far as we shall carry the theory of the $\{u_n\}$-, $\{v_n\}$-sets. We now point out how the results of the present section can be applied to the convergence question in methods of best approximation.

THEOREM 21. *Let $u_n(x)=\Phi_n(x)$, $v_n(x)=\Omega_n(x)$, $(n=0,1,\cdots)$ be two basic sets relative to a method $\mathfrak{M}$ of best approximation, and let $\Phi_n(x)$, $\Omega_n(x)$ be analytic in a region $\mathfrak{S}$ and continuous in $\mathfrak{S}+C$, C being the boundary of $\mathfrak{S}$. We further assume:*

(i) *Conditions A and B hold.*

(ii) *If*[20] *$\{h_n(x)\}$ is any sequence of functions analytic in $\mathfrak{S}$ and continuous in $\mathfrak{S}+C$, then the operators $\{L_i\}$ that define the method $\mathfrak{M}$ are term-wise applicable to every series $\Sigma c_n h_n(x)$ that converges uniformly on every closed point set in $\mathfrak{S}$.*

[20] We actually make use of (ii) only for the sequences $h_n(x)=\Phi_n(x)$, $\Omega_n(x)$. The following observation is of interest: By the Corollary to Condition A, each $\Omega_n(x)$ has a $\Phi_n(x)$-expansion uniformly convergent on any closed point set in $\mathfrak{S}$. Therefore ((ii) and Theorem 2)

$$(14)\qquad \Omega_n(x)=\Phi_n(x)+c_{n,n+1}\Phi_{n+1}(x)+c_{n,n+2}\Phi_{n+2}(x)+\cdots.$$

From (14) it is seen that Condition B will certainly be fulfilled if the coefficients $c_{n,n+i}$ are chosen sufficiently small.

Under these conditions, if $f(x)$ is any function analytic in $\mathfrak{A}$ and continuous in $\mathfrak{A} + C$, the approximating "polynomials" $s_n(x)$ of $f(x)$ (relative to the basic set $\{\Omega_n(x)\}$) converge uniformly to $f(x)$ on every closed point set in $\mathfrak{A}$.

For: By Theorem 8, $f(x)$ has an $\Omega_n(x)$-expansion, uniformly convergent on any closed set in $\mathfrak{A}$: (a) $f(x) = \sum_0^\infty F_n\Omega_n(x)$. On the other hand, if $s_n(x)$ is the best "polynomial" of n-th order, then (using (6) with Φ_n replaced by Ω_n),

(b) $$s_0(x) = f_0\Omega_0(x), \qquad s_n(x) - s_{n-1}(x) = f_n\Omega_n(x),$$

(c) $$f_n = L_n[s_n(x) - s_{n-1}(x)].$$

Also ((2)),

(d) $$L_i[s_i(x)] = L_i[f(x)].$$

The theorem will be established if we show that $F_n = f_n$, $(n = 0, 1, \cdots)$, since (cf. (8)) $s_n(x) = \sum_{i=0}^n f_i\Omega_i(x)$. By (ii) we may operate [30] with L_i on (a):

(e) $L_i[f] = F_0L_i[\Omega_0] + F_1L_i[\Omega_1] + \cdots + F_iL_i[\Omega_i]$, $(i = 0, 1, \cdots)$.

Taking $i = 0$: $L_0[F] = F_0L_0[\Omega_0]$; $L_0[s_0] = F_0L_0[\Omega_0]$; therefore $f_0L_0[\Omega_0] = F_0L_0[\Omega_0]$, and $f_0 = F_0$ since $L_0[\Omega_0] = 1$. Now assume that $F_r = f_r$, $(r = 0, 1, \cdots, i-1)$; we shall complete the induction for i: (e) reduces to

$$L_i[f] = f_0L_i[\Omega_0] + f_1L_i[\Omega_1] + \cdots + f_{i-1}L_i[\Omega_{i-1}] + F_i;$$
$$L_i[s_i] = L_i[s_0] + L_i[s_1 - s_0] + \cdots + L_i[s_{i-1} - s_{i-2}] + F_i;$$
$$L_i[s_i] = L_i[s_{i-1}] + F_i; \qquad L_i[s_i - s_{i-1}] = F_i;$$

therefore $f_i = F_i$. This completes the proof.[31]

A comparison of (8) with (28) and (39) yields the

COROLLARY. *The coefficients in the Ω_n-expansion for $f(x)$ are given (variously) by*

$$(55)\quad f_n = \frac{1}{2\pi i}\int_C g(t)L_n(t)\,dt = \frac{1}{2\pi i}\int_C f(t)M_n(t)\,dt = L_n[s_n(x) - s_{n-1}(x)].$$

PENNSYLVANIA STATE COLLEGE.

---

[30] We also use (5): $L_i[\Omega_n] = 0$, $i < n$.

[31] It is worth noting that in this theorem we do not assume the uniqueness Condition C. A possible choice of $v_n(x)$ is $v_n(x) \equiv u_n(x) = \Phi_n(x)$. Hence, if we omit the details that make Theorem 21 precise, the *sense* of the theorem is contained in the statement: *If $\{\Phi_n\}$, $\{\Omega_n\}$ are two basic sets "sufficiently close" to each other, they give essentially the same convergence properties to the respective best approximating "polynomials."*

GROUPS CONTAINING FIVE AND ONLY FIVE SQUARES.

By G. A. Miller.

Let G represent a group such that the squares of its operators are five and only five distinct operators including the identity. When these five squares constitute a group it results that there are two and only two such groups which are not direct products. This is a special case of a general theorem relating to groups whose squares constitute a cyclic subgroup.[1] When these squares do not constitute a group there are three possible cases, as follows: Two of them are of order 4 and two of order 2, two are of order 3 and two of order 2, or four of them are of order 2. In each of these cases the identity constitutes the fifth square. This is always found among the squares since it is the square of itself.

When two of the squares are of order 4 then the squares generate the abelian group of order 8 and of type $(2, 1)$ which includes two operators of order 4 and one of order 2 which are non-squares. These 8 operators constitute an invariant subgroup of G which corresponds to an abelian quotient group of order 2^m and of type $(1, 1, 1, \cdots)$. This invariant subgroup appears in an invariant abelian subgroup of order 16 and of type $(3, 1)$. Since G includes operators of order 4 whose common square is not equal to the square of the operators of order 4 which appear in this subgroup of order 16 the latter operators are not commutative with the former and they give rise to commutators of order 4 with respect to the operators of order 8 in the given subgroup of type $(3, 1)$. As these automorphisms are of order 2 these commutators are transformed into their inverses under G.

For the sake of brevity in the statements it will be assumed in what follows that G is not the direct product of a group containing five and only five operators which are squares and of an abelian group of order 2^m and of type $(1, 1, 1, \cdots)$. The order of G cannot be less than 32 and when it is of this order its central is the four group contained in the given invariant abelian subgroup of type $(3, 1)$. There is one such G in which each of the operators of this invariant subgroup is transformed into its inverse and all of the remaining operators are of order 4 and have a common square which is distinct from the square of the operators of order 4 contained in the given

[1] G. A. Miller, *Proceedings of the National Academy of Sciences*, vol. 20 (1934), pp. 203-206.

subgroup of order 16. There is no such G in which each of the operators of this subgroup is transformed into its third power but there are two such G's in which the commutator subgroup is the cyclic group of order 4 which is not generated by an operator of order 8 contained in G. In one of these two groups 8 of the additional operators are of order 2 while in the other all of the additional operators are of order 4 but have two distinct squares. Hence there results the following theorem: *Every group which has the property that it contains five and only five operators which are squares, including such an operator of order* 4, *involves at least one of the three groups of order* 32 *which have this property.*

To determine all the possible groups of order 64 which contain five and only five operators which are squares including one of order 4 it is therefore only necessary to extend each of the three groups noted in the preceding theorem by 32 additional operators. These include 16 operators which are commutative with an operator of order 4 which is a square and hence each possible set of 32 additional operators includes an operator of order 2 which is commutative with this operator of order 4. To the first of the three given groups of order 32 we can adjoin three such sets of 32 operators and thus obtain three G's of order 64. In two of these the given added operator of order 2 has only two conjugates under G while in the third it has four such conjugates. To each of the other two given groups of order 32 we can adjoin only one such set of 32 operators. As all of these groups are distinct, there results the theorem that *there are five and only five groups of order* 64 *which separately have the property that they involve five and only five squares including such an operator of order* 4.

If such a G of order 128 exists it contains a subgroup of order 64 composed of all of its operators which are commutative with an operator t_1 of order 4 which is a square under G and all of whose operators of order 4 have a common square. Suppose first that this subgroup is abelian. An operator t_2 of order 4 in G whose square is different from t_1^2 is then commutative with at least 8 of the operators of this subgroup of order 64 and all of these operators besides the identity are of order 2. Hence G involves an operator of order 2 which is not contained in the subgroup generated by its squares but is commutative with all of its operators. It is therefore a direct product of an abelian group of order 2^m and of type $(1, 1, 1, \cdots)$ and of a group which involves five and only five operators which are squares thereunder. Since such direct products have been excluded it results that the given subgroup of order 64 cannot be abelian.

Its commutator subgroup cannot include an operator of order 4 since

such an operator would be either t_1 or t_1^3 and hence it could not arise from an operator of order 4 or from an operator of order 2 contained in this subgroup. The operators of these two orders contained in this subgroup therefore generate a characteristic subgroup of order 32 under G. As each of the co-sets of this subgroup with respect to the subgroup formed by the squares under G involves 4 operators of order 2 and such an operator of order 2 cannot be transformed under this subgroup into itself multiplied by t_2^2 it results that the commutator subgroup of this group of order 64 is generated by t_1^2. Its central is of order 16 and either of type $(2, 1, 1)$ or of type $(3, 1)$. In the former case it is easy to verify that no G can exist while in the latter case there is one such G. Hence *there are nine groups which have the property that each of them contains five and only five operators which are squares thereunder including at least one of order* 4. *Three of these are of order* 32, *five are of order* 64, *and one is of order* 128.

Suppose that the five operators which are squares under G include an operator of order 3 and hence two such operators. Since all the operators which are squares under G are relatively commutative [2] it results that such a G involves two and only two operators of order 2 which are squares and that each of the operators of order 4 in G transforms its operators of order 3 into their inverses. Hence such a G contains a subgroup of index 2 which is the direct product of its subgroup of order 3 and an abelian group of order 2^m and of type $(1, 1, 1, \cdots)$. Each of its remaining operators is of order 4 since its operators of order 4 have two and only two distinct squares and every two operators of order 4 in G which have distinct squares are non-commutative. It therefore results that the commutator subgroup of G is the cyclic group of order 6 and that *there is one and only one group which satisfies the condition that it has five and only five operators which are squares including an operator of order* 3. *The order of this group is* 48 *and it contains the direct product of the group of order* 3 *and the abelian group of order* 8 *and of type* $(1, 1, 1)$.

It remains to consider the possible cases when the five operators which are squares include exactly four (s_1, s_2, s_3, s_4) which are of order 2. These four operators generate an abelian group whose order is either 8 or 16. We shall first prove that this order cannot be 16. If s_1, s_2, s_3, s_4 would generate a group H of order 16 then H would appear in the central of G for reasons which follow. Such an operator s_1 could not be non-commutative with another operator s_5 of order 2 contained in G for if it were s_1 and s_5 would generate

[2] *Ibid.*, vol. 19 (1933), pp. 1054-1057.

the octic group which would involve two of the three operators s_2, s_3, s_4 since the conjugate of an operator which is a square has the same property. As s_1 and these two operators would generate the four group the four operators s_1, s_2, s_3, s_4 could not then generate a group of order 16.

If one of these four operators s_2 would be non-commutative with an operator t_1 of order 4 contained in G, it may be assumed that $t_1^2 = s_1$ and that t_1 transforms s_2 and s_3 into each other and is therefore commutative with s_2s_3. The group of order 8 generated by s_1, s_2, s_3 would be invariant under t_1 and $(t_1s_2)^2$ would equal $s_1s_2s_3$, which is impossible if s_1, s_2, s_3, s_4 generate a group of order 16. It therefore follows that if H were of order 16 it would appear in the central of G and every two operators of order 4 contained in G which have different squares would be non-commutative. If such a G exists we may assume without loss of generality that $t_1^2 = s_1$, $t_2^2 = s_2$, $t_3^2 = s_3$, and $t_4^2 = s_4$. It results directly that t_2 transforms t_1 into itself multiplied by one of the following five operators s_1, s_2, s_1s_2, $s_1s_2s_3$, $s_1s_2s_4$ since H includes the commutator subgroup of G. We shall first prove that the fourth of these cases is impossible and hence the fifth is also impossible.

In the fourth case t_1 and t_2 together with H generate a group of order 64 and t_1t_2 may be assumed to be t_3. The operator t_4 transforms each of the operators t_1, t_2, t_1t_2 into itself multiplied respectively by one of the following operators: s_1, s_4, s_1s_4, $s_1s_4s_2$, $s_1s_4s_3$; s_2, s_4, s_2s_4, $s_2s_4s_1$, $s_2s_4s_3$; s_3, s_4, s_3s_4, $s_3s_4s_1$, $s_3s_4s_2$. This is impossible because t_4 transforms t_1t_2 into itself multiplied by the product of its two commutators with t_1 and t_2. It therefore results that t_2 transforms t_1 into itself multiplied by one of the following three operators; s_1, s_2, s_1s_2. If we can prove that the first of these is impossible it will also prove that the second is impossible. Hence we assume that the first condition is satisfied until we arrive at a contradiction.

The group of order 64 generated by H, t_1, t_2 then involves only operators of order 4 in addition to H. The operator t_3 gives rise to the following commutators with respect to t_1, t_2, t_1t_2 respectively: s_1, s_3, s_1s_3; s_2, s_3, s_2s_3; s_2, s_3, s_2s_3. Hence there is only one such subgroup of order 128 possible. In this t_3 gives rise to the following commutators s_3, s_2, s_2s_3 with respect to t_1,t_2, t_1t_2 respectively. Hence t_4 gives rise with respect to t_1, t_2, t_1t_2, t_3, t_1t_3, t_2t_3, $t_1t_2t_3$ respectively to the following commutators: s_1, s_4, s_1s_4; s_2, s_4, s_2s_4; s_2, s_4, s_2s_4; s_3, s_4, s_3s_4; s_1, s_4, s_1s_4; s_3, s_4, s_3s_4; 1, s_1s_4, s_2s_4, s_3s_4. As these are obviously inconsistent such a subgroup of order 128 cannot appear in G. The existence of such a G therefore implies that t_2 gives rise to the commutators s_1s_2 with respect to t_1 and hence it gives rise to the commutators s_2s_3, s_3s_4 with respect to t_3 and t_4 respectively. As this is impossible, since

it would give rise to too many squares, it has been proved that *when a group involves five and only five operators which are squares and four of them are of order 2 then these four operators generate an invariant subgroup of order* 8.

If an operator of order 2 in G would not be commutative with every operator of this invariant subgroup H then it would be non-commutative with one of its operators s_1 which is a square and it and s_1 would generate an octic group involving three operators which are squares. This operator would therefore be commutative with exactly half of the operators of H since it is commutative with the remaining square of order 2. Hence at least one of the squares of G would be invariant under G since the co-sets with respect to H are invariant and therefore two of these squares would have this property. An operator whose square is a non-invariant operator of G would therefore be commutative with every operator of H and hence all the operators of the co-set with respect to H to which it belongs have the same square. This square is therefore invariant under G, which is contrary to the hypothesis. That is, we arrived at a contradiction by assuming that an operator of order 2 in G is not commutative with every operator of H. If an operator of order 4 in G were non-commutative with a square it would transform exactly two squares among themselves. Since no operator could be non-commutative with all of the four squares these squares could not be transformed under G according to a group of degree 4. It has been noted that only two squares could not be non-invariant. Hence, it results that *H is in the central of G.*

It is easy to see that H is the central of G since this central cannot contain an operator of order 4 and if it would contain an operator of order 2 which does not appear in H then G would be a direct product. Each of the possible groups appears in one and only one of the following three categories: The first is composed of those in which the product of every two distinct squares of order 2 is a non-square, the second of those in which at least two operators of order 4 which have different squares are commutative, the third of those in which the product of two squares is a square but no two operators of order 4 which have different squares are commutative. In the first case it may be assumed that the four squares of order 2 are s_1, s_2, s_3, $s_1s_2s_3$. In each of the other two cases it may be assumed that the squares of order 2 are s_1, s_2, s_1s_2, s_3. The smallest order of a group which satisfies the conditions under consideration is 64. We proceed to determine all the groups of this order which belong to these three categories in the given order.

We shall first prove that each of these groups contains a definite subgroup of order 32. To prove this we first extend H by an operator t_1 of order 4 whose square is s_1 and thus obtain an abelian subgroup of type $(2, 1, 1)$.

This subgroup is then extended by t_2 whose square is s_2 so as to obtain a subgroup of order 32. We proceed to prove that this can always be so selected that t_1t_2 is of order 2 and hence it is completely determined. If t_1t_2 were of order 4 its square may be assumed to be different from s_3. Hence t_3 whose square is s_3 would have to transform this subgroup of order 32 so as to give rise to four commutators including at least one of the form s_1s_2. An operator of order 4 with respect to which t_3 gives rise to this commutator and t_3 have a product of order 2 since the square of such a product could not be of the form s_1s_2. This proves the following theorem: *If a group involves five and only five operators which are squares including four of order* 2 *and if the product of no two distinct ones of these operators of order* 2 *is a square then the group contains a subgroup of order* 32 *generated by these squares and two operators of order* 4 *having distinct squares and a product of order* 2.

It may now be assumed that all the groups of order 64 belonging to the first of the three categories under consideration contain this subgroup of order 32 generated by H, t_1, t_2 where t_1t_2 is of order 2. There is obviously one and only one such group in which the product of every two operators of order 4 which have different squares is of order 2. There is also one and only one such group in which t_1t_3 is of order 2 but t_2t_3 is not of this order. To prove that in each of the remaining groups of this category each of the additional operators is of order 4 it is only necessary to note that $t_1t_2t_3$ could not be of order 2 in some one of them. This results from the fact that we may assume that t_2t_3 is not of order 2 since we would otherwise get a group which is conjugate with the one already considered, and hence no commutator of the form s_1s_2 could arise from t_3. Each of these remaining groups therefore contains three pairs of operators of order 4 whose products are of order 2 and which are distinct modulo H and have distinct squares. In particular, each of these remaining groups involves three conjugate subgroups of order 32 which have the abelian subgroup of type $(1, 1, 1, 1)$ in common.

Since an operator whose square is $s_1s_2s_3$ appears among those which are added to the given group of order 32 it may be assumed without loss of generality that t_3 transforms t_1 into itself multiplied by s_2 and that it transforms t_1t_2 into itself multiplied by one of the following four operators: 1, s_1s_2, s_1s_3, s_2s_3. Hence there are two additional such groups of order 64. In one of these the commutator subgroup is of order 4 while in the other it is of order 8. It therefore results that *there are four and only four groups of order* 64 *which separately satisfy the condition that they contain five and only five operators which are squares, including four of order* 2, *and that the product of no two squares is a square.*

The second category of groups under consideration contains by hypothesis the abelian group of type $(2, 2, 1)$. To extend this so as to obtain a G of order 64 it is necessary to add thereto an operator t_4 of order 4 whose square s_4 is the fourth square of order 2 in G. Since t_4 is not commutative with any operator of order 4 contained in this subgroup it must give rise to four distinct commutators with respect thereto and hence it transforms into its inverse at least one of these operators of order 4. Since H contains three subgroups of order 4 which include the square of the operator of order 4 which is transformed into its inverse by t_4 and one of these subgroups corresponds to two possible G's there results the following theorem: *Four and only four groups of order* 64 *have the property that each of them contains five operators which are squares thereunder and contains the abelian group of type* $(2, 2)$.

It remains to determine the groups of order 64 which separately satisfy the condition that no two of their operators of order 4 which have different squares are commutative but that the product of the squares of two such operators is one of the four squares of order 2 contained therein. We may assume that t_1 and t_2 are non-commutative and that $t_4^2 = s_1s_2$. We shall first consider the case when t_1t_2 is of order 2 and extend the subgroup of order 32 generated by H, t_1, t_2 by t_4 so as to obtain one of the groups of order 64 which satisfies the given condition. It is easy to verify that t_4 cannot transform one of the three operators t_1, t_2, t_1t_2 into itself multiplied by s_3. It can also not transform t_1t_2 into itself multiplied by one of the following operators: s_1s_2, s_1s_3, 1, s_1, s_2 since its product with one of the three operators t_1, t_2, t_1t_2 has s_3 for *its* square. If t_1t_2 is transformed into itself multiplied by s_1s_2 the group is completely determined. The commutator subgroup of this group is of order 4.

When t_1t_2 is transformed into itself multiplied by $s_1s_2s_3$ the commutator subgroup is of order 8. There is one and only one such group and hence there are two possible groups of order 64 in which t_1t_2 is of order 2. When t_1t_2 is of order 4 the square of their product may be one of the two operators s_1, s_2 or it may be s_3. In the former case we may assume witnout loss of generality that the square of t_1t_2 is s_2. We do not need to consider the case when t_4 transforms t_1t_2 into itself multiplied by s_1 since G would then contain two operators of order 4 having different squares whose product would be of order 2. As before t_4 could not transform one of the operators t_1, t_2, t_1t_2 into itself multiplied by s_3. Moreover, t_4 could not transform t_1t_2 into itself multiplied by one of the following operators: 1, s_2, $s_1s_2s_3$, s_2s_3. If it transforms it into itself multiplied by s_1s_2 then G is completely determined and involves only

operators of order 4 in addition to H. Hence *there are seven and only seven groups of order* 64 *which separately satisfy the conditions that each contains five and only five operators which are squares, four being of order* 2, *and that no two operators of order* 4 *which have different squares are commutative but the product of two squares is a square.*

There is no upper limit for the orders of the remaining possible groups since every such group can be extended so as to obtain a group whose order is four times the order of the given group provided this group contains a subgroup of index 2 such that each of the remaining operators is of order 4. It is easy to verify that this condition is satisfied by groups in each of the three given categories and that when it is satisfied we may use the direct product of this group and a group of order 2 and adjoin to it an operator of order 4 which is commutative with an operator of order 4 not contained in the given subgroup of index 2, has the same square as the latter operator, and transforms this subgroup in the same manner as the given operator of order 4 transforms it. The resulting group can then be used in the same way to construct such a group of four times its own order and this process can be repeated indefinitely. Each of the infinite systems of groups thus obtained contains exactly five operators which are squares thereunder and four of these operators are of order 2.

CORRECTION AND ADDITION TO "COMPLEMENTS OF POTENTIAL THEORY."[1]

By GRIFFITH C. EVANS.

Dr. F. G. Dressel has called my attention by means of an example to the necessity of a correction for Lemma II, p. 217, in the above mentioned memoir. In fact, the theorem of Daniell, quoted in the lemma, does not apply. The lemma should read as follows:

LEMMA II. *Let $f(x)$ be bounded and measurable in the Borel sense, $g_1(x)$, $g_2(x)$ of bounded variation and $g_1(x)$ or $g_2(x)$ continuous, $a \leq x \leq b$; then $g(x) = g_1(x)g_2(x)$ is of bounded variation, and*

$$\int_a^b f(x)dg(x) = \int_a^b f(x)g_1(x)dg_2(x) + \int_a^b f(x)g_2(x)dg_1(x). \tag{1.1}$$

In the proof of the lemma omit (c), line 7, p. 218, replacing it by (d), and replace line 22, p. 218, by the inequality

$$t_m(b) - t_m(a) \leq t(b) - t(a),$$

for the total variation functions. Let $g_2(x)$, say, be continuous. There is then no need of $g_{2m}(x)$, and the proof, in the case of $f(x)$ continuous, ends with line 3, p. 219. The extension to $f(x)$, bounded and measurable in the Borel sense, is the same as before.

We note also the following:

LEMMA II′. *Let $g_1(x)$, $g_2(x)$ be of bounded variation, $a \leq x \leq b$, t_1 the total variation of g_1, and N_2 the upper bound of $| g_2(x) |$ over $a \leq x \leq b$. Then*

$$\left| \int_a^b g_1(x)dg_2(x) \right| \leq | g(b) - g(a) | + t_1 N_2.$$

[1] *American Journal of Mathematics*, vol. 54 (1932), pp. 213-234. The example given by Dr. Dressel is the following:

$$\begin{aligned} f(x) &= 1, && 0 \leq x \leq 2, \\ g_1(x) = g_2(x) &= 0, && 0 \leq x \leq 1, \\ &= 1, && 1 < x \leq 2, \end{aligned}$$

for which the left-hand member of (1.1) has the value 1 and the right-hand member the value 0.

In fact, let $g_{1n}(x)$ be the continuous polygonal approximation to $g_1(x)$, of Lemma II. Then, $g_{1n}(b)g_2(b) - g_{1n}(a)g_2(a) = g(b) - g(a)$, and

$$\int_a^b g_{1n}(x)dg_2(x) = g(b) - g(a) - \int_a^b g_2(x)dg_{1n}(x)$$

$$\left|\int_a^b g_{1n}(x)dg_2(x)\right| \leq |g(b) - g(a)| + \{t_{1n}(b) - t_{1n}(a)\}N_2$$
$$\leq |g(b) - g(a)| + t_1N_2.$$

But

$$\lim_{n=\infty}\int_a^b g_{1n}(x)dg_2(x) = \int_a^b g_1(x)dg_2(x),$$

so that the inequality is established.

In the proof of Lemma IV, p. 222, Lemma II′ should be cited for the inequality of line 11, p. 224, without making use of the previous equation.

A general theorem of the type of Lemma II is the following one.

THEOREM. *Let $f(x)$ be bounded and measurable Borel, and $g_1(x)$, $g_2(x)$ of bounded variation, $a \leq x \leq b$; and let e_1, e_2 be the sets respectively of values of x corresponding to the points of discontinuity of $g_1(x)$, $g_2(x)$. Then* (1. 1) *is valid provided e_1 and e_2 have no points in common.*

There is evidently no loss in generality in assuming $g_1(x), g_2(x)$ to be not negative and monotone-increasing. The function $g(x)$ will then be of the same sort. Let $f(x)$ be continuous, and write

$$g_i(x) = \alpha_i(x) + \beta_i(x), \qquad (i = 1, 2),$$

where $\alpha_i(x)$ is continuous and $\beta_i(x)$ is the corresponding "function of discontinuities." We have

$$\int_a^b f(x)dg(x) = \int_a^b fd(\alpha_1\alpha_2) + \int_a^b fd(\alpha_1\beta_2) + \int_a^b fd(\alpha_2\beta_1) + \int_a^b fd(\beta_1\beta_2),$$

and by Lemma II, the identity (1. 1) may be applied to every integral except the last. By proving that it applies to the last integral, the identity will be established for $f(x)$ continuous, and may then be extended to $f(x)$ bounded and measurable Borel, as before.

Accordingly it remains to prove that

$$\int_a^b fd(\beta_1\beta_2) = \int_a^b f\beta_1 d\beta_2 + \int_a^b f\beta_2 d\beta_1,$$

assuming $f(x)$ to be continuous.

Consider first the case where β_1 and β_2 are step functions with merely a finite number of jumps B_i, C_j at values $x = b_i$, $x = c_j$ respectively, with $b_i \neq c_j$ for all i, j. Then, evidently,

$$\int_a^b fd(\beta_1\beta_2) = \Sigma_i f(b_i)\beta_2(b_i)B_i + \Sigma_j f(c_j)\beta_1(c_j)C_j$$
$$= \int_a^b f(x)\beta_2(x)d\beta_1(x) + \int_a^b f(x)\beta_1(x)d\beta_2(x),$$

where the integrals of the right-hand member are general (Daniell) integrals.

Let now $\beta_1(x)$ and $\beta_2(x)$ be arbitrary functions of discontinuities, so that they may have a denumerable infinity of finite jumps. We define $\beta_{1n}(x)$ as a step function, approximating to $\beta_1(x)$, with merely a finite number of jumps. In fact, let $b_1, b_2, \cdots, b_{k_n}$ be the values of x at which the jump of $\beta_1(x)$ is $\geqq 1/n$. It may happen that a or b is a b_i. The function $\beta_{1n}(x)$ is to have discontinuities only at $b_1, \cdots, b_{n_k}$ and at these points it is to have the same discontinuities as $\beta_1(x)$, viz.,

$$\beta_{1n}(a) = \beta_1(a)$$
$$\beta_{1n}(b_i) - \beta_{1n}(b_i \pm 0) = \beta_1(b_i) - \beta_1(b_i \pm 0).$$

It is clear that

$$\text{(I)} \qquad \lim_{n=\infty} \beta_{1n}(x) = \beta_1(x), \qquad a \leq x \leq b.$$

For on the one hand, $\beta_{1n}(x) \leq \beta_1(x)$. And on the other hand, given $\epsilon > 0$, we can find n so that the sum of all the finite jumps which are each in value $< 1/n$, will be $< \epsilon$. Consequently for n sufficiently large, $\beta_{1n}(x) > \beta_1(x) - \epsilon$, $a \leq x \leq b$. Moreover, since $\beta_{1n}(x_2) - \beta_{1n}(x_1)$ is merely the sum of discontinuities of $\beta_1(x)$ belonging to those finite jumps of $\beta_1(x)$, in the interval $x_1 \leq x \leq x_2$, each of which is in value $\geq 1/n$, we have

$$\text{(II)} \qquad \beta_{1n}(x_2) - \beta_{1n}(x_1) \leq \beta_1(x_2) - \beta_1(x_1), \qquad a \leq x_1 < x_2 \leq b.$$

We define similarly functions $\beta_{2m}(x)$ approximating to $\beta_2(x)$, and similar properties (I), (II) hold for the functions $\beta_{mn}(x) = \beta_{1n}(x) \cdot \beta_{2m}(x)$. In fact,

$$\begin{aligned}
&\beta_{1n}(x_2)\beta_{2m}(x_2) - \beta_{1n}(x_1)\beta_{2m}(x_1) \\
&\quad = \{\beta_{1n}(x_2) - \beta_{1n}(x_1)\}\beta_{2m}(x_2) + \beta_{1n}(x_1)\{\beta_{2m}(x_2) - \beta_{2m}(x_1)\} \\
&\qquad \leq \{\beta_1(x_2) - \beta_1(x_1)\}\beta_2(x_2) + \beta_1(x_1)\{\beta_2(x_2) - \beta_2(x_1)\} \\
&\quad = \beta_1(x_2)\beta_2(x_2) - \beta_1(x_1)\beta_2(x_1).
\end{aligned}$$

The hypotheses (I), (II) are the hypotheses of Daniell's theorem,[2] whence

$$\lim_{m=\infty}\int_a^b f(x)\beta_{1n}(x)\,d\beta_{2m}(x)=\int_a^b f(x)\beta_{1n}(x)\,d\beta_2(x).$$

But also

$$\lim_{n=\infty}\int_a^b f(x)\beta_{1n}(x)\,d\beta_2(x)=\int_a^b f(x)\beta_1(x)\,d\beta_2(x),$$

as an elementary property of the general integral. Since e_1, e_2 have no common elements, there are no common points of discontinuity of β_{1n}, β_{2m}, and

$$\int_a^b f\,d\beta_{mn}=\int_a^b f\beta_{1n}\,d\beta_{2m}+\int_a^b f\beta_{2m}\,d\beta_{1n}.$$

By successive passages to the limit, the identity is established.

University of California,
Berkeley, California.

[2] P. J. Daniell, "Further properties of the general integral," *Annals of Mathematics*, vol. 21 (1920), pp. 203-220. See p. 218.

ON A CERTAIN CLASS OF ORTHOGONAL POLYNOMIALS.

By A. Tartler.

Introduction. Let $\psi(x)$ denote a bounded non-decreasing function—"characteristic function"—with infinitely many points of increase on the finite or infinite interval (a, b) and such that

all "moments" $\alpha_i = \int_a^b x^i d\psi(x)$ exist $(i = 0, 1, \cdots)$, with $\alpha_0 > 0$.[1]

The object of this paper is to study the system of orthogonal and normal polynomials

$$(1) \quad u_n(x; d\dot{\psi}) \equiv u_n(x) \equiv \bar{a}_n(x^n - \bar{s}_n x^{n-1} + \cdots) \equiv \bar{a}_n U_n(x; d\dot{\psi}) \equiv \bar{a}_n U_n(x) \qquad (n = 0, 1, \cdots; \bar{a}_n > 0)$$

corresponding to the more general characteristic function of bounded variation

$$\dot{\psi}(x) = \int_a^x (t - \alpha) d\psi(t) \quad (a < \alpha < b), \text{ with moments } \beta_i = \int_a^b x^i d\dot{\psi}(x) \qquad (i = 1, 2, \cdots),$$

and their relation to the system of polynomials

$$(2) \quad \begin{aligned} \phi_n(x) \equiv \phi_n(x; d\psi) &\equiv a_n x^n + a_{n,n-1} x^{n-1} + \cdots \quad (a_n > 0;\ n = 0, 1, \cdots) \\ &\equiv a_n \Phi_n(x) \equiv a_n(x^n - S_n x^{n-1} + \cdots), \end{aligned}$$

having the fundamental property

$$(3) \qquad \int_a^b \phi_m(x) \phi_n(x) d\psi(x) = \delta_{m,n} \qquad (m, n = 0, 1, \cdots).$$

(3) is equivalent to

$$(4) \qquad \int_a^b \phi_n(x) G_{n-1}(x) d\psi(x) = 0,$$

where $G_s(x) = \sum_{i=0}^{s} g_i x^i$ here and hereafter stands for an arbitrary polynomial

[1] We assume the non-existence of numbers c, d such that

$$\int_a^c d\psi(x) = \int_d^b d\psi(x) = 0, \qquad (a < c,\ d < b).$$

of degree $\leqq s$. Our main purpose is to show how far the known properties of the system (2) are extendable to (1). The results obtained are an extension of those announced, without proof, by J. Shohat.[2]

1. *Some needed properties of orthogonal polynomials.*

$$(5)\quad \text{(i)}\ \begin{cases} \Phi_n(x) = (x - c_n)\Phi_{n-1}(x) - \lambda_n\Phi_{n-2}(x)\,;\ \lambda_n = a^2{}_{n-2}/a^2{}_{n-1}\,; \\ c_n = S_n - S_{n-1} = \int_a^b x\phi^2{}_{n-1}(x)\,d\psi(x) \qquad (n \geqq 2)\,; \\ c_1 = S_1 = \alpha_1/\alpha_0,\quad \lambda_1 = \alpha_0\,;\quad s_n = \sum_{i=1}^{n} c_i \qquad (n \geqq 1). \end{cases}$$

(ii) The polynomials $\Phi_n(x)$ are the denominators of the successive convergents of the "associated" continued fraction[3]

$$(6)\qquad \int_a^b \frac{d\psi(y)}{x-y} \sim \frac{\lambda_1\,/}{/\,x - c_1} - \frac{\lambda_2\,/}{/\,x - c_2} - \cdots.$$

(iii) If $x_{1,n}, \cdots, x_{n,n}$ denote the zeros of $\Phi_n(x)$, then

$$(7)\qquad a < x_{1,n+1} < x_{1,n} < x_{2,n+1} < x_{2,n} < \cdots < x_{n,n} < x_{n+1,n+1} < b.$$

(iv) Darboux' formula,[4] which are of fundamental importance in the discussion which follows

$$(8)\quad \begin{cases} K_n(x, t\,; d\psi) \equiv K_n(x, t) \equiv \sum_{i=0}^{n} \phi_i(x)\phi_i(t) \\ \qquad = \dfrac{a_n}{a_{n+1}}\,\dfrac{\phi_{n+1}(x)\phi_n(t) - \phi_{n+1}(t)\phi_n(x)}{x-t} \\ K_n(x, x\,; d\psi) \equiv K_n(x) \equiv \sum_{i=0}^{n} \phi_i{}^2(x) \\ \qquad = \dfrac{a_n}{a_{n+1}}\,[\phi'_{n+1}(x)\phi_n(x) - \phi_{n+1}(x)\phi'_n(x)]. \end{cases}$$

(v) Let $\{\eta_{i,n}\}$, $\{\xi_{i,n}\}$ $(i = 1, 2, \cdots, n)$ denote respectively the zeros of $\phi_n(x\,; d\psi_1)$, $\phi_n(x\,; d\psi_2)$, with

$$\psi_1(x) \equiv \int_a^x (t-a)\,d\psi(t),\quad \psi_2(x) \equiv \int_a^x (b-t)\,d\psi(t)\quad (a \leqq x \leqq b):$$

$$(9)\quad a < x_{1,n+1} < \xi_{1,n} < x_{1,n} < \eta_{1,n} < x_{2,n+1} < \xi_{2,n} < x_{2,n} < \eta_{2,n} < x_{3,n+1} < \cdots < x_{n,n+1} < \xi_{n,n} < x_{n,n} < \eta_{n,n} < x_{n+1,n+1} < b.$$

---

[2] Jacques Chokhate (J. Shohat), "Sur les fractions continues algébriques," *Comptes Rendus*, vol. 191 (1930), p. 474.

[3] O. Perron, *Die Lehre von den Kettenbrüchen*, Teubner, 1913, p. 377.

[4] Darboux, "Mémoire sur l'approximation des fonctions de très grands nombres," *Journal de Mathématiques* (3), vol. 4 (1878), pp. 5-56, 377-416.

2. *Existence of the system of orthogonal polynomials $U_n(x; d\psi)$.* The fundamental problem is to derive conditions assuring the existence of a sequence of polynomials $U_n(x)$ defined in (1), satisfying either one of the equivalent conditions of orthogonality:

$$\text{(10)} \quad \begin{cases} \int_a^b U_n(x) G_{n-1}(x) d\psi(x) = \int_a^b U_n(x) G_{n-1}(x)(x-\alpha) d\psi(x) = 0 \\ \hfill (n = 1, 2, \cdots), \\ \int_a^b U_n(x) U_m(x)(x-\alpha) d\psi(x) = 0 \quad (n \neq m;\ n, m = 0, 1, 2, \cdots). \end{cases}$$

Using (3, 10), we get, writing $(x-\alpha) U_n(x) = \sum_{i=0}^{n+1} A_i \phi_i(x)$:

$$\text{(11)} \qquad (x-\alpha) U_n(x) = A_{n+1}\phi_{n+1}(x) + A_n \phi_n(x).$$
$$A_{n+1} = \frac{1}{a_{n+1}}, \qquad A_n \phi_n(\alpha) = -\frac{\phi_{n+1}(\alpha)}{a_{n+1}}.$$

If $\phi_n(\alpha) \neq 0$, A_n is uniquely determined. If $\phi_n(\alpha) = 0$, then necessarily $\phi_{n+1}(\alpha) \neq 0$ (see (7)) and A_n does not exist. This, combined with Darboux' formula (8), leads to

THEOREM I. *A necessary and sufficient condition that $U_n(x)$ satisfying (10) exist for a given n, is: $\phi_n(\alpha) \neq 0$. $U_n(x)$ is then uniquely determined:*

$$\text{(12)} \quad U_n(x) \equiv \frac{1}{a_{n+1}\phi_n(\alpha)} \frac{\phi_{n+1}(x)\phi_n(\alpha) - \phi_{n+1}(\alpha)\phi_n(x)}{x-\alpha} \equiv \frac{K_n(x, \alpha)}{a_n \phi_n(\alpha)}.$$

COROLLARY. *If $U_n(x)$ exist, then* $U_n(\alpha) = \dfrac{K_n(\alpha)}{a_n\phi_n(\alpha)} \neq 0.$

If $\phi_{n+1}(\alpha) = 0$, then by (11) $A_n = 0$ and

$$\text{(13), (13')} \qquad U_n(x) \equiv \frac{\phi_{n+1}(x)}{a_{n+1}(x-\alpha)}; \qquad \int_a^b U_n(x) G_n(x) d\psi(x) = 0;$$

i.e., the degree of the arbitrary polynomial $G_n(x)$ being here as high as n — the degree of $U_n(x)$. Conversely, by (12, 3), if a polynomial $U_n(x) \equiv x^n + \cdots$ satisfies (13′), then necessarily $\phi_{n+1}(\alpha) = 0$. As an immediate consequence of Theorem I we state the important

THEOREM II. *$\phi_n(\alpha) \neq 0$ for $n \geqq 1$[5] implies the existence of a set of orthogonal polynomials $U_n(x) = x^n + \cdots$ of all degrees $(n = 0, 1, \cdots)$ satisfying (10) and uniquely determined by means of (12).*

Hereafter we assume $\phi_n(\alpha) \neq 0$ $(n \geqq 1)$ unless explicitly stated otherwise.

[5] Infinitely many such α exist in any subinterval of (a, b).

3. *Normalization of the system $U_n(x)$.* By virtue of (8, 3)

$$\int_a^b K_n(x, \alpha) G_n(x) d\psi(x) = -\frac{\phi_{n+1}(\alpha) g_n}{a_{n+1}}.$$

Take here $G_n(x) \equiv K_n(x, \alpha)$ and use (12):

$$\int_a^b K_n^2(x, \alpha) d\psi(x) = -\frac{a_n}{a_{n+1}} \phi_{n+1}(\alpha)\phi_n(\alpha),$$

$$(14) \qquad \int_a^b U_n^2(x) d\psi(x) = -\frac{\phi_{n+1}(\alpha)}{a_n a_{n+1} \phi_n(\alpha)} \equiv \frac{\rho_n}{\bar{a}_n^2} \qquad (\rho_n = \pm 1;\ \bar{a}_n > 0),$$

$$\frac{1}{\bar{a}_n^2} \equiv \left| \frac{\phi_{n+1}(\alpha)}{\phi_n(\alpha)} \right| \frac{1}{a_n a_{n+1}}$$

$$(15) \qquad \int_a^b u_n^2(x) d\psi(x) = \rho_n = \operatorname{sgn}\left[-\frac{\phi_{n+1}(\alpha)}{\phi_n(\alpha)}\right].$$

Thus the integral in (15) is positive or negative, contrary to what we shall call the "ordinary" case, i. e., that of a *monotonic* characteristic function. Turning to (7), we get at once

$$(16) \qquad \begin{aligned} \rho_n &= +1 \quad \text{for} \quad \alpha < x_{1,n+1}, \quad \text{or} \quad x_{k,n} < \alpha < x_{k+1,n+1}; \\ \rho_n &= -1 \quad \text{for} \quad x_{k,n+1} < \alpha < x_{k,n}, \quad \text{or} \quad x_{n+1,n+1} < \alpha. \end{aligned}$$

4. *The recurrence relation for $U_n(x)$.* Write

$$(17) \qquad U_n(x) = (x - \bar{c}_n) U_{n-1}(x) + P_{n-2}(x) \qquad (\bar{c}_n = \text{const.}),$$

where $P_{n-2}(x)$ is a polynomial of degree $\leqq n - 2$. Making use of (10), we get at once:

$$(18) \qquad 0 = \int_a^b P_{n-2}(x) G_{n-3}(x)(x - \alpha) d\psi(x).$$

The degree of $P_{n-2}(x)$ cannot be less than $n - 3$. For otherwise, we could take in (18) $G_{n-3}(x) \equiv (x - \alpha) P_{n-2}(x)$ and thus render the integrand non-negative. $P_{n-2}(x)$ cannot be of degree $n - 3$ for then (18) would be equivalent to (13′), which in turn implies $\phi_{n-2}(\alpha) = 0$, contrary to our assumption (§ 2). *Hence, $P_{n-2}(x)$ is actually of degree $n - 2$.* Moreover, (18) being nothing but the condition of orthogonality (10), $P_{n-2}(x)$ differs from $U_{n-2}(x)$ by a constant factor only, so that (17) becomes

$$(19) \qquad U_n(x) = (x - \bar{c}_n) U_{n-1}(x) - \bar{\lambda}_n U_{n-2}(x) \qquad (n \geqq 2;\ \bar{c}_n,\ \bar{\lambda}_n = \text{const.}).$$

We thus obtain for $\{U_n(x)\}$ a recurrence relation precisely of the same typ[e] as (5). (19) yields through (10, 14) by comparing coefficients:

$$(20)\quad \begin{cases} \bar{\lambda}_n = \dfrac{\int_a^b U_{n-1}(x)x^{n-1}d\psi(x)}{\int_a^b U_{n-2}(x)x^{n-2}d\psi(x)} = \rho_{n-1}\rho_{n-2}\dfrac{\bar{a}^2_{n-2}}{\bar{a}^2_{n-1}} & (n \geqq 2), \\ \bar{c}_n = \dfrac{\int_a^b xU^2_{n-1}(x)\,d\psi(x)}{\int_a^b U^2_{n-1}(x)\,d\psi(x)} = \rho_{n-1}\int_a^b xu^2_{n-1}(x)\,d\psi(x) & (n \geqq 2;\ \bar{c}_1 = \beta_1/\beta_0), \\ \bar{s}_n = \sum_{i=1}^{n} \bar{c}_i, \qquad \bar{c}_n = \bar{s}_n - \bar{s}_{n-1}. \end{cases}$$

It follows that *in the case under consideration* $\bar{\lambda}_n$ *are not all positive,* contrary to the ordinary case.

Introduce, as in the ordinary case (Perron, *l. c.*), the "associated" power series and continued fraction

$$(21)\quad \int_a^b \frac{d\psi(y)}{x-y} \sim \sum_{i=0}^{\infty} \frac{\beta_i}{x^{i+1}} \sim \frac{\bar{\lambda}_1/}{/q_1(x)} - \frac{\bar{\lambda}_2/}{/q_2(x)} - \cdots \quad (q_i(x)\text{-polynomials}).$$

The n-th convergent of the latter we denote by $P_n(x)/Q_n(x)$, $Q_n(x)$ being of degree μ_n $(n \geqq 0)$. Then, its fundamental property

$$(22)\qquad \int_a^b \frac{d\psi(y)}{x-y} - \frac{P_n(x)}{Q_n(x)} = \left(\frac{1}{x^{\mu_n+\mu_{n+1}}}\right)^{6}$$

leads to the orthogonality property for $Q_n(x)$:

$$\int_a^b Q_n(x)\,G_{\mu_n-1}(x)\,d\psi(x) = 0.$$

Hence, we may identify $Q_n(x)$ with $U_{\mu_n}(x)$, or with $U_{\mu_n-1}(x)$ according as $\phi_{\mu_n}(\alpha)$ is, or is not, zero. If $\phi_n(\alpha) \neq 0$ for $n \geqq 1$, the degrees of the denominators of the successive convergents in (21) differ by one and all the $q_n(x)$ are of the first degree, as in (6) above:

$$(23)\quad \int_a^b \frac{d\psi(y)}{x-y} \sim \frac{\bar{\lambda}_1/}{/x-\bar{c}_1} - \frac{\bar{\lambda}_2}{/x-\bar{c}_2} - \cdots \quad (\bar{\lambda}_2, \cdots, \bar{c}_1, \cdots \text{ from (20)}).$$

We take $\bar{\lambda}_1 = \beta_0$, for, by (22),

$$\int_a^b \frac{d\psi(y)}{x-y} - \frac{\bar{\lambda}_1}{x-\bar{c}_1} = \left(\frac{1}{x^3}\right).$$

5. *The zeros of* $U_n(x)$ *compared with those of* $\phi_{n+1}(x)$. Denote the zeros of $U_n(x)$ by $\bar{x}_{i,n}$ $(i = 1, 2, \cdots, n;\ n \geqq 1)$. By (12):

---

[6] $(1/x^s)$ generally stands for $c_1/x^s + c_2/x^{s+1} + \cdots$ $(c_1 \neq 0)$.

$$(24)\quad (x_{i,n+1}-\alpha)(x_{i+1,n+1}-\alpha)U_n(x_{i,n+1})U_n(x_{i+1,n+1})$$
$$=\frac{\phi^2_{n+1}(\alpha)}{a^2_{n+1}\phi_n{}^2(\alpha)}\phi_n(x_{i,n+1})\phi_n(x_{i+1,n+1})<0.$$

Considering the sign of the product of the first two factors in the left-hand member of (24), we readily arrive at

THEOREM III. *The interval* $(x_{i,n+1}, x_{i+1,n+1})$ *contains either no zeros, or one zero, of* $U_n(x)$, *according as* α *is, or is not, an interior point of it.*

Remark. (13) shows that if α is one of the zeros of $\phi_{n+1}(x)$, its remaining n zeros are precisely those of $U_n(x)$. This case was excluded and is mentioned here merely as a limiting case when α tends to a zero of $\phi_{n+1}(x)$ (Cf. § 6).

COROLLARY. *If* $\alpha < x_{1,n+1}$ *or* $> x_{n+1,n+1}$, *the zeros of* $U_n(x)$ *separate those of* $\phi_{n+1}(x)$.

We proceed to investigate more closely the case when α is an interior point of one of the intervals $(x_{k,n+1}, x_{k+1,n+1})$ $(k=1,2,\cdots,n)$. Since $x_{k,n+1}<x_{k,n}<x_{k+1,n+1}$, it is convenient to consider two cases:

(i) $x_{k,n+1}<\alpha<x_{k,n}$. In (12) put $x=x_{n+1,n+1}$:

$$(x_{n+1,n+1}-\alpha)U_n(x_{n+1,n+1})=-\frac{1}{a_{n+1}}\frac{\phi_{n+1}(\alpha)}{\phi_n(\alpha)}\phi_n(x_{n+1,n+1}).$$

Here $x_{n+1,n+1}-\alpha>0$, $-\dfrac{\phi_{n+1}(\alpha)}{\phi_n(\alpha)}<0$ (see (7)'), $\phi_n(x_{n+1,n+1})>0$; hence $U_n(x_{n+1,n+1})<0$ and $U_n(x)$ *has one and only one zero in the interval* $(x_{n+1,n+1}, +\infty)$.

(ii) $x_{k,n}<\alpha<x_{k+1,n+1}$. We find in a similar manner that $U_n(x)$ *has one and only one zero in the interval* $(-\infty, x_{1n+1})$.

We proceed further to specify the values of α for which the zeros of $U_n(x)$, for a given n, include either a or b (assumed to be finite). To this end consider $\Phi_n(x)\equiv\Phi_n(x;d\psi_2)$ for which

$$\int_a^b\left\{\frac{(b-x)\Phi_n(x)}{x-\xi_{k,n}}\right\}G_{n-1}(x)(x-\xi_{k,n})d\psi(x)=0\qquad(1\leqq k\leqq n).$$

Hence, by virtue of Theorem I (uniqueness):

$$U_n(x;(x-\xi_{k,n})d\psi)\equiv-\frac{(b-x)\Phi_n(x;d\psi_2)}{x-\xi_{k,n}}.$$

Similarly we treat the point $x=a$ by means of $\Phi_n(x; d\psi_1)$. In other words, if α is a zero of the polynomial $\Phi_n(x; d\psi_{1,2})$, one of the zeros of $U_n(x)$ coincides correspondingly with a or b. This conclusion fully harmonizes with the inequalities and the results of §§ 2, 3.

6. *The zeros* $\{\bar{x}_{i,n}\}$ *of* $U_n(x)$ *as functions of* α.

THEOREM IV. $\{\bar{x}_{i,n}\}$ *increase with* α.

Proof. Differentiate with respect to α the relation $K_n(\bar{x}_{i,n}, \alpha)=0$ (see (12)):

$$\frac{d\bar{x}_{i,n}}{d\alpha} = -\frac{\dfrac{\partial K_n(\bar{x}_{i,n}, \alpha)}{\partial\alpha}}{\dfrac{\partial K_n(\bar{x}_{i,n}, \alpha)}{\partial\bar{x}_{i,n}}} \qquad (i=1,2,\cdots,n;\ n\geqq 1). \tag{25}$$

Develop the right-hand member in (25), making use of (12):

$$\frac{d\bar{x}_{i,n}}{d\alpha} = \frac{g(\alpha, \bar{x}_{i,n})}{g(\bar{x}_{i,n}, \alpha)} \qquad (g(x,y)\equiv\phi'_{n+1}(x)\phi_n(y)-\phi_{n+1}(y)\phi'_n(x)).$$

The desired result, namely

$$\frac{d\bar{x}_{i,n}}{d\alpha} > 0 \qquad (i=1,2,\cdots,n;\ n\geqq 1)$$

will be established if we succeed in showing that

$$g(\alpha, \bar{x}_{i,n})g(\bar{x}_{i,n}, \alpha) > 0.$$

But this latter inequality follows from the readily verifiable identity

$$g(x,y)g(y,x)\equiv g(x)g(y)$$
$$+[\phi'_n(x)\phi'_{n+1}(y)-\phi'_n(y)\phi'_{n+1}(x)][\phi_{n+1}(x)\phi_n(y)-\phi_{n+1}(y)\phi_n(x)],$$
$$\left(g(x)\equiv g(x,x)=\frac{a_{n+1}}{a_n}\sum_{i=0}^{n}\phi_i^2(x)\right)$$

which leads to

$$g(\alpha, \bar{x}_{i,n})g(\bar{x}_{i,n}, \alpha)=g(\alpha)g(\bar{x}_{i,n})=\frac{a^2_{n+1}}{a_n^2}\sum_{i=0}^{n}\phi_i^2(\alpha)\sum_{i=0}^{n}\phi_i^2(\bar{x}_{i,n})>0.$$

Remark. The above general theorem holds for any real α, inside or outside (a,b).

The results of § 5, together with Theorem IV, are sufficient to describe completely the behavior of $\{\bar{x}_{i,n}\}$ as α varies increasingly from $-\infty$ to $+\infty$. This description is summarized in the following table. It will be recalled

(§ 2) that when α is a zero of $\phi_n(x)$, $U_n(x)$ does not exist, but $U_{n-1}(x)$ necessarily exists; at this point it is convenient to regard it as the polynomial $U_n(x)$, with one of its zeros infinite.

α	$\bar{x}_{i,n}$
$\alpha < x_{1,n+1}$	$x_{i,n+1} < \bar{x}_{i,n} < x_{i+1,n+1} \quad (i=1,2,\cdots,n)$
* $\alpha = x_{1,n+1}$	$\bar{x}_{i,n} = x_{i+1,n+1} \quad (i=1,2,\cdots,n)$
$x_{1,n+1} < \alpha < \xi_{1,n}$	$x_{i,n+1} < \bar{x}_{i-1,n} < x_{i+1,n+1} \ (i=2,3,\cdots,n+1;\ x_{n+2,n+1}\equiv b)$.
$\alpha = \xi_{1,n}$	$\bar{x}_{i,n} = \xi_{i+1,n} \ (i=1,2,\cdots,n-1), \quad \bar{x}_{n,n} = b$
$\xi_{1,n} < \alpha < x_{1,n}$	$x_{i,n+1} < \bar{x}_{i-1,n} < x_{i+1,n+1} \ (i=2,3,\cdots,n), \quad \bar{x}_{n,n} > b$
$\alpha = x_{1,n}$	$\bar{x}_{n,n} = +\infty$, or $\bar{x}_{1,n} = -\infty$
$x_{1,n} < \alpha < \eta_{1,n}$	$\bar{x}_{1,n} < a,\ x_{i,n+1} < \bar{x}_{i,n} < x_{i+1,n+1} \quad (i=2,3,\cdots,n)$
$\alpha = \eta_{1,n}$	$\bar{x}_{1,n} = a,\ \bar{x}_{i,n} = \eta_{i,n} \quad (i=2,3,\cdots,n)$
$\eta_{1,n} < \alpha < x_{2,n+1}$	$a < \bar{x}_{1,n} < x_{1,n+1},\ x_{i,n+1} < \bar{x}_{i,n} < x_{i+1,n+1} \ (i=2,3,\cdots,n)$
** $\alpha = x_{2,n+1}$	$\bar{x}_{1,n} = a,\ \bar{x}_{i,n} = \eta_{i,n} \quad (i=2,3,\cdots,n)$
α varies from $x_{2,n+1}$ to $x_{n+1,n+1}$	$\bar{x}_{i,n}$ varies as from * to **, with proper changes of indices.
$x_{n+1,n+1} < \alpha$	$x_{i,n+1} < \bar{x}_{i,n} < x_{i+1,n+1} \quad (i=1,2,\cdots,n)$.

7. *On the separation of the zeros of $U_n(x)$ and $U_{n+1}(x)$.* Here we use $K_n(x,\alpha)$ instead of $U_n(x)$. Consider first the case when $a < \alpha < x_{1n+1}$. Here (§ 5) the zeros $\bar{x}_{i,n}$ of $K_n(x,\alpha)$ separate the zeros of $\phi_{n+1}(x)$:

$$(26) \qquad a < x_{1,n+1} < \bar{x}_{1,n} < x_{2,n+1} < \bar{x}_{2n} < \cdots < \bar{x}_{n,n} < x_{n+1,n+1} < b.$$

We note that if (26) holds for $n = n_0$, it does so for $n < n_0$, for the hypothesis $\alpha < x_{1,n+1}$ implies $\alpha < x_{1,m}$, $m < n+1$, by virtue of (7). Furthermore,

$$(27) \qquad K_{n+1}(\bar{x}_{i,n},\alpha)K_{n+1}(\bar{x}_{i+1,n},\alpha) = \phi^2{}_{n+1}(\alpha)\phi_{n+1}(\bar{x}_{i,n})\phi_{n+1}(\bar{x}_{i+1,n}),$$

$$(28) \qquad K_{n+1}(x_{i,n+1},\alpha)K_{n+1}(x_{i+1,n+1},\alpha) = K_n(x_{i,n+1},\alpha)K_n(x_{i+1,n+1},\alpha).$$

The right-hand member of (27) being negative by virtue of (26), it follows that $K_{n+1}(x,\alpha)$ changes sign an odd number of times in each of the intervals $(\bar{x}_{i,n}, \bar{x}_{i+1,n})$ $(i=1,2,\cdots,n-1)$. Moreover, the right-hand member of (28) being also negative, we conclude that it changes sign in each of these intervals only once. Hence,

$$\alpha < x_{1,n+1} \text{ implies } \bar{x}_{1,n+1} < \bar{x}_{1,n} < \bar{x}_{2,n+1} < \bar{x}_{2,n} < \cdots < \bar{x}_{n,n} < \bar{x}_{n+1,n+1}.$$

The same inequalities hold if $x_{n+1,n+1} < \alpha$.

Assume now that α separates two of the zeros of $\phi_{n+1}(x)$; say, $x_{k,n+1} < \alpha < x_{k+1,n+1}$. Here (§ 5) $K_n(x,\alpha)$ has no zero in $(x_{k,n+1}, x_{k+1,n+1})$ and has one zero in each of the remaining intervals $(x_{i,n+1}, x_{i+1,n+1})$ $(i \neq k)$. If in (27) $i \neq k-1$, its right-hand member is negative, and $K_{n+1}(x,\alpha)$ changes sign in the corresponding intervals $(\bar{x}_{i,n}, \bar{x}_{i+1,n})$ $(i \neq k)$ at least once. We now assert that $K_{n+1}(x,\alpha)$ changes sign twice in the interval $(\bar{x}_{k-1,n}, \bar{x}_{k,n})$; more precisely, it changes sign in each of the subintervals $(\bar{x}_{k-1,n}, x_{k,n+1})$, $(x_{k+1,n+1}, \bar{x}_{k,n})$. In fact,

$$K_{n+1}(\bar{x}_{k-1,n}, \alpha) = \phi_{n+1}(\bar{x}_{k-1,n})\phi_{n+1}(\alpha),$$

$$K_{n+1}(x_{k,n+1}, \alpha) = K_n(x_{k,n+1}, \alpha) = -\frac{a_n}{a_{n+1}} \frac{\phi_{n+1}(\alpha)\phi_n(x_{k,n+1})}{x_{k,n+1} - \alpha},$$

$$K_{n+1}(\bar{x}_{k-1,n}, \alpha) K_{n+1}(x_{k,n+1}, \alpha) = -\frac{a_n}{a_{n+1}} \frac{\phi^2_{n+1}(\alpha)\phi_{n+1}(\bar{x}_{k-1,n})\phi_n(x_{k,n+1})}{x_{k,n+1} - \alpha}.$$

Furthermore, since $\operatorname{sgn} \phi_{n+1}(\bar{x}_{k-1,n}) = (-1)^{n+k}$, $\operatorname{sgn} \phi_n(x_{k,n+1}) = (-1)^{n+k-1}$,

$$K_{n+1}(\bar{x}_{k-1,n}, \alpha) K_{n+1}(x_{k,n+1}, \alpha) < 0,$$

and similarly,

$$K_{n+1}(\bar{x}_{k,n}, \alpha) K_{n+1}(x_{k+1,n+1}, \alpha) < 0.$$

The last two inequalities prove our assertion. Thus we state

THEOREM V. (i) *$\alpha < x_{1,n+1}$ or $> x_{n+1,n+1}$ implies: the zeros of $U_n(x)$ separate those of $U_{n+1}(x)$;* (ii) *$x_{k,n+1} < \alpha < x_{k+1,n+1}$ implies: each of the intervals $(\bar{x}_{i,n}, \bar{x}_{i+1,n})$ $(i \neq k-1)$ contains one zero and the interval $(\bar{x}_{k-1,n}, \bar{x}_{k,n})$ contains two zeros of $U_{n+1}(x)$.*

(The one remaining zero of $U_{n+1}(x)$ is either $< \bar{x}_{1,n}$ or $> \bar{x}_{n,n}$).

It is known for the ordinary case that the zeros of $\Phi_n(x)$ for n very large are everywhere dense in (a, b), provided

$$(29) \qquad \int_{a_1}^{b_1} d\psi(x) \neq 0 \qquad (a \leqq a_1 < b_1 \leqq b)$$

This, combined with Theorem V leads to the

COROLLARY. *Under (29) the zeros of $U_n(x)$ for n very large are everywhere dense in (a, b).*

8. *The mechanical quadratures formula related to $U_n(x)$.* We consider the mechanical quadratures formula—a direct application of the Lagrange interpolation formula—

$$(30) \qquad \int_a^b G_{n-1}(x)\,d\psi(x) = \sum_{i=1}^{n} h_{i,n} G_{n-1}(\zeta_i)$$

$$\left(h_{i,n} \equiv \int_a^b \frac{\prod_{i=1}^{n}(x-\zeta_i)\,d\psi(x)}{(x-\zeta_i)\left\{\prod_{i=1}^{n}(x-\zeta_i)\right\}'_{x=\zeta_i}} \right),$$

where $\{\zeta_i\}$ denote n distinct points arbitrarily chosen. If we write

$$G_{2n-1}(x) = G_{n-1}(x) \prod_{i=1}^{n}(x-\zeta_i) + G^{(1)}_{n-1}(x)$$

and make use of the orthogonality properties (10), we arrive at

THEOREM VI. *The mechanical quadratures formula* (30) *holds for* $G_{2n-1}(x)$ *if, and only if, the points* ζ_i *are zeros of* $U_n(x)$.

We thus get a formula of Gauss' type

$$(31) \qquad \int_a^b G_{2n-1}(x)\,d\psi(x) = \sum_{i=1}^{n} \bar{H}_{i,n} G_{2n-1}(\bar{x}_{i,n})$$

$$\left(\bar{H}_{i,n} \equiv \int_a^b \frac{U_n(x)\,d\psi(x)}{(x-\bar{x}_{i,n})U'_n(\bar{x}_{i,n})} \right).$$

We get further, knowing that $\bar{x}_{i,n} \neq \alpha$, and taking in (31) successively

$$G_{2n-1}(x) \equiv \left\{ \frac{U_n(x)}{(x-\bar{x}_{i,n})U'_n(\bar{x}_{i,n})} \right\}^2, \quad \left\{ \frac{U_n(x)}{(x-\bar{x}_{i,n})U'_n(\bar{x}_{i,n})} \right\}^2 (x-\alpha):$$

$$\bar{H}_{i,n} = \int_a^b \left\{ \frac{U_n(x)}{(x-\bar{x}_{i,n})U'_n(\bar{x}_{i,n})} \right\}^2 (x-\alpha)\,d\psi(x),$$

$$(32) \qquad \bar{H}_{i,n} = \frac{1}{\bar{x}_{i,n}-\alpha} \int_a^b \left\{ \frac{U_n(x)}{(x-\bar{x}_{i,n})U'_n(\bar{x}_{i,n})} \right\}^2 (x-\alpha)^2\,d\psi(x)$$

$$\begin{matrix} <0 \\ >0 \end{matrix} \quad \text{for } \bar{x}_{i,n} \begin{matrix} <\alpha \\ >\alpha \end{matrix}.$$

Here again we find an essential difference between the case under consideration and the ordinary one (where all the coefficients in the mechanical quadratures formula of type (31) are positive).

We proceed to derive an interesting expression for $\bar{H}_{i,n}$ in terms of $K_n(x)$. The orthogonality property (10) rewritten as

$$\int_a^b \left\{ \frac{U_n(x)(x-\alpha)}{x-\bar{x}_{i,n}} \right\} G_{n-1}(x)(x-\bar{x}_{i,n})\,d\psi(x) = 0$$

shows the existence of a polynomial of degree n

$$U_n(x;\,(x-\bar{x}_{i,n})\,d\psi)\equiv\frac{U_n(x;\,d\psi)\,(x-\alpha)}{x-\bar{x}_{i,n}},$$

orthogonal in (a,b) with respect to the characteristic function

$$\psi_{i,n}\equiv\int_a^x (t-\bar{x}_{i,n})\,d\psi(t).$$

We derive successively:

$$U_n(x;\,d\psi)=\frac{U_n(x;\,d\psi_{i,n})\,(x-\bar{x}_{i,n})}{x-\alpha};\quad U'_n(\bar{x}_{i,n};\,d\psi)=\frac{U_n(\bar{x}_{i,n};\,d\psi_{i,n})}{\bar{x}_{i,n}-\alpha}.$$

Substitute in (31) and apply (12, 8, 3):

$$\bar{H}_{i,n}=\frac{\bar{x}_{i,n}-\alpha}{K_n(\bar{x}_{i,n};\,d\psi)}\qquad (i=1,2,\cdots,n). \tag{33}$$

(33) is another proof of the inequalities in (32). It also may give indication[7] as to the asymptotic behavior of $\bar{H}_{i,n}$ for $n\to\infty$.

9. *Extension of Darboux' formulae.* The recurrence relation (19) readily leads to

$$\sqrt{\rho_{n-1}\rho_n\bar{\lambda}_{n+1}}\,\frac{u_n(x)u_{n-1}(y)-u_n(y)u_{n-1}(x)}{x-y}=u_{n-1}(x)u_{n-1}(y)$$
$$+\sqrt{\rho_{n-2}\rho_{n-1}\bar{\lambda}_n}\,\frac{u_{n-1}(x)u_{n-2}(y)-u_{n-1}(y)u_{n-2}(x)}{x-y},$$

$$\bar{K}_n(x,y)\equiv\sum_{i=0}^{n}u_i(x)u_i(y)=\frac{\bar{a}_n}{\bar{a}_{n+1}}\,\frac{u_{n+1}(x)u_n(y)-u_{n+1}(y)u_n(x)}{x-y},$$

$$\bar{K}_n(x,x)\equiv\bar{K}_n(x)=\frac{\bar{a}_n}{\bar{a}_{n+1}}\,[u'_{n+1}(x)u_n(x)-u_{n+1}(x)u'_n(x)].$$

Thus Darboux' formulae (8) *hold in our case without any modification.*

10. *A mechanical quadratures formula with a fixed interior point.* Consider the mechanical quadratures formula

$$\int_a^b G_n(x)\,d\psi(x)=\sum_{i=0}^{n}H_{i,n}G_n(\zeta_i) \tag{34}$$
$$\left[H_{i,n}\equiv\int_a^b\frac{\prod_{i=0}^{n}(x-\zeta_i)\,d\psi(x)}{(x-\zeta_i)\left\{\prod_{i=0}^{n}(x-\zeta_i)\right\}'_{x=\zeta_i}}\right],$$

where the points $\{\zeta_i\}$ $(i=0,1,\cdots,n)$ are distinct and $\zeta_0=\alpha$ is arbitrarily fixed inside (a,b). We may show by the method of § 8 that (34) holds for $G_{2n}(x)$, provided $\phi_n(\alpha)\neq 0$ $(n=1,2,\cdots)$, $\zeta_{i,n}=\bar{x}_{i,n}$ $(i=1,2,\cdots,n)$, the zeros of $U_n(x)$, so that (see (33))

[7] This could be illustrated by means of Hermite polynomials.

$$(35)\quad H_{i,n}=\frac{\bar H_{i,n}}{\bar x_{i,n}-\alpha}=\frac{1}{K_n(\bar x_{i,n})}\ (i=1,2,\cdots,n),\qquad H_{0,n}=\frac{1}{K_n(\alpha)}.$$

Assume (a,b) *to be finite.* Since $\bar x_{i,n}$ $(i=2,3,\cdots,n-1)$ and at least one of the two zeros $\bar x_{1,n}$, $\bar x_{n,n}$ are always in (a,b), (35) shows that *the corresponding* $\bar H_{i,n}$ *tend to zero as* $n\to\infty$, *in all cases for which it is known that* $H_{i,n}\to 0$. (See § 11 below).

We get further, taking in (34) successively

$$G_{2n}(x)\equiv\left\{\frac{(x-\alpha)U_n(x)}{(x-\bar x_{i,n})(\bar x_{i,n}-\alpha)U'_n(\bar x_{i,n})}\right\}^2,\quad\left\{\frac{U_n(x)}{U_n(\alpha)}\right\}^2,\quad 1:$$

$$H_{i,n}=\int_a^b\left\{\frac{(x-\alpha)U_n(x)}{(x-\bar x_{i,n})(\bar x_{i,n}-\alpha)U'_n(\bar x_{i,n})}\right\}^2 d\psi(x)\qquad(i=1,2,\cdots,n),$$

$$H_{0,n}=\int_a^b\left\{\frac{U_n(x)}{U_n(\alpha)}\right\}^2 d\psi(x),$$

$$(36)\qquad \sum_{i=0}^{n}H_{i,n}=\int_a^b d\psi(x).$$

We see that here *all* $H_{i,n}$ *are positive.*

11. *Tchebycheff inequalities related to* $U_n(x)$. Denote the set of points α, $\{\bar x_{i,n}\}$ $(i=1,2,\cdots,n)$ by $y_{1,n+1}<y_{2,n+1}<\cdots<y_{n+1,n+1}$, change the numbering of $H_{i,n}$ in (34) accordingly and rewrite it as

$$\int_a^b G_{2n}(x)d\psi(x)=\sum_{i=1}^{n+1}H_iG_{2n}(y_{i,n+1}).$$

Following Stieltjes [8] (and Markoff), construct $G_{2n}(x)$ subject to the following conditions:

$$\begin{aligned}G_{2n}(y_{i,n+1})&=1 &&(i=1,2,\cdots,k),\\ G_{2n}(y_{i,n+1})&=0 &&(i=k+1,k+2,\cdots,n+1),\\ G'_{2n}(y_{i,n+1})&=0 &&(1\leqq i\leqq n+1,\ i\neq k).\end{aligned}$$

These $2n+1$ conditions determine $G_{2n}(x)$ uniquely. Moreover, $G'_{2n}(x)$ has n zeros at the points $y_{i,n+1}$ $(i=1,2,\cdots,k-1,k+1,\cdots,n+1)$ and, by Rolle's theorem, $k-1$ zeros inside $(y_{i,n+1},y_{i+1,n+1})$ $(i=1,2,\cdots,k-1)$, and $n-k$ zeros inside $(y_{i,n+1},y_{i+1,n+1})$ $(i=k,\ k+1,\cdots,n+1)$, with $n+(k-1)+n-k=2n-1$. It follows readily that

$$G_{2n}(x)\geqq 0\ \text{for all}\ x,\ \geqq 1\ \text{for}\ x\leqq y_{k,n+1},$$

[8] Stieltjes, "Quelques recherches sur les quadratures dites mécaniques," *Œuvres*, vol. 1, pp. 377-396.

$$\int_a^b G_{2n}(x)\,d\psi(x) = H_1 + H_2 + \cdots + H_k \geqq \int_a^{y_{k,n+1}} G_{2n}(x)\,d\psi(x)$$
$$\begin{cases} k = 1, 2, \cdots, n+1, & \text{if } \bar{x}_{n,n} \leqq b \\ k = 1, 2, \cdots, n, & \text{if } \bar{x}_{n,n} > b \end{cases}$$

$$H_1 + H_2 + \cdots + H_k \geqq \int_a^{y_{k,n+1}} d\psi(x) \tag{37}$$
$$\begin{cases} k = 1, 2, \cdots, n+1, & \text{if } \bar{x}_{n,n} \leqq b \\ k = 1, 2, \cdots, n, & \text{if } \bar{x}_{n,n} > b. \end{cases}$$

Similarly, using the polynomial $T_{2n}(x)$ such that

$$\begin{aligned} T_{2n}(y_{i,n+1}) &= 0 \qquad (i = 1, 2, \cdots, k); \\ T_{2n}(y_{i,n+1}) &= 1 \qquad (i = k+1, k+2, \cdots, n+1); \\ T'_{2n}(y_{i,n+1}) &= 0 \qquad (1 \leqq i \leqq n+1;\ i \neq k+1): \end{aligned}$$

$$H_{k+1} + H_{k+2} + \cdots + H_{n+1} \geqq \int_{y_{k+1,n+1}}^{b} d\psi(x) \qquad (k = 1, 2, \cdots, n), \tag{38}$$

and combining this with (36):

$$H_1 + H_2 + \cdots + H_k \leqq \int_a^{y_{k+1,n+1}} d\psi(x) \qquad (k = 1, 2, \cdots, n). \tag{39}$$

The inequalities (37, 39) constitute an extension to our case of the important Tchebycheff inequalities. It follows readily that

$$H_k \leqq \int_{y_{k-1,n+1}}^{y_{k+1,n+1}} d\psi(x) \qquad (k = 2, 3, \cdots, n), \tag{40}$$
$$H_1 \leqq \int_a^{y_{2,n+1}} d\psi(x), \qquad H_{n+1} \leqq \int_{y_{n,n+1}}^{b} d\psi(x).$$

Hence, if $\psi(x)$ is continuous in the finite interval (a, b) which contains no subinterval (a_1, b_1) such that $\int_{a_1}^{b_1} d\psi(x) = 0$, then (see Corollary to Theorem V) $H_i \to 0$ as $n \to \infty$ $(i = 1, 2, \cdots, n+1)$. By virtue of (35) we infer that $K_n(\alpha) \to \infty$ as $n \to \infty$ for any fixed α. *This result combined with a theorem due to Hamburger*[9] *gives a direct and elementary proof of the important fact that the moment problem for a finite interval is determined.*

12. *On the associated continued fraction.*

THEOREM VI. *If $r(x)$ is a continuous function having s changes of sign between a and b and $\psi(x)$ is of the nature indicated, then in the associated continued fraction*

$$F(x) \equiv \int_a^b \frac{r(y)\,d\psi(y)}{x - y} \sim \frac{\lambda_1/}{/q_1(x)} - \frac{\lambda_2/}{/q_2(x)} - \cdots \tag{41}$$

[9] H. Hamburger, "Über eine Erweiterung des Stieltjesschen Momentenproblem," *Mathematische Annalen*, vol. 81 (1920), pp. 235-319, Theorem XVII.

the degrees of the polynomials $q_i(x)$ $(i=1,2,\cdots)$ *cannot exceed* $s+1$.

Proof. We have formally, denoting the i-th convergent of (41) by $\Omega_i(x)/\Phi_i(x)$, with $\Phi_i(x)\equiv x^{\mu_i}+p_{\mu_i-1}x^{\mu_i-1}+\cdots$:

$$F(x)=\sum_{n=0}^{\infty}\frac{\alpha_n}{x^{n+1}}\ \left(\alpha_n\equiv\int_a^b r(y)y^n d\psi(y)\right),\quad F(x)\Phi_i(x)-\Omega_i(x)=\left(\frac{1}{x^{\mu_{i+1}}}\right),$$

and expanding the left-hand member:

$$\alpha_j p_0+\alpha_{j+1}p_1+\cdots+\alpha_{j+\mu_i}=0 \qquad (j=0,1,\cdots,\mu_{i+1}-2),$$

which is equivalent to

$$\int_a^b r(x)\Phi_i(x)G_{\mu_{i+1}-2}(x)d\psi(x)=0 \qquad (i=1,2,\cdots). \tag{42}$$

Were a certain $q_i(x)$ in (41) of degree $>s+1$, we would have

$$\mu_{i+1}-\mu_i>s+1,\qquad \mu_{i+1}-2\geqq\mu_i+s,$$

and we could render (42) impossible by choosing $G_{\mu_{i+1}-2}(x)$ in (42) so that $r(x)G_{\mu_{i+1}-2}(x)\geqq 0$ for $a\leqq x\leqq b$. The results of § 13 below show that *the upper bound for the degree of* $q_i(x)$ *as given in Theorem VI is the best possible.*

13. *The case* $\phi_n(\alpha)=0$. Here (§ 2) $U_n(x)$ does not exist, $U_{n-1}(x)$ and $U_{n+1}(x)$, however, necessarily exist (since $\phi_{n-1}(\alpha)\phi_{n+1}(\alpha)\neq 0$). Moreover, by (13′),

$$\int_a^b U_{n-1}(x)G_{n-1}(x)(x-\alpha)d\psi(x)=0$$

We get, writing

$$U_{n+1}(x)=(x^2+\bar{b}_{n+1}x+\bar{c}_{n+1})U_{n-1}(x)+P_{n-2}(x) \tag{43}$$

$(P_{n-2}(x)$ — polynomial of degree $\leqq n-2)$:

$$\int_a^b P_{n-2}(x)G_{n-3}(x)(x-\alpha)d\psi(x)=0 \qquad \text{(see (10))}. \tag{44}$$

(44) can be satisfied in the following cases only

(i) $P_{n-2}(x)\equiv 0$; (ii) $P_{n-2}(x)\equiv -l_{n+1}U_{n-2}(x)$, if $\phi_{n-2}(\alpha)\neq 0$;

(iii) $P_{n-2}(x)\equiv -l_{n+1}U_{n-3}(x)$, if $\phi_{n-2}(\alpha)=\phi_n(\alpha)=0$.

(i) *is impossible.* In fact, it leads, through (13′, 10), to

$$U_{n+1}(x)=(x^2+\bar{b}_{n+1}x+\bar{c}_{n+1})U_{n-1}(x),\qquad \int_a^b x^n U_{n-1}(x)(x-\alpha)d\psi(x)=0;$$

while on the other hand by (13),

$$\int_a^b x^n U_{n-1}(x)(x-\alpha)\,d\psi(x) = \int_a^b x^n \frac{\phi_n(x)}{a_n(x-\alpha)}(x-\alpha)\,d\psi(x) = \frac{1}{a_n^2} \neq 0.$$

(ii) Here (43) becomes

$$(45) \qquad U_{n+1}(x) = (x^2 + \bar{b}_{n+1}x + \bar{c}_{n+1})U_{n-1}(x) - l_{n+1}U_{n-2}(x).$$

We find as above (see (20)):

$$l_{n+1} = \frac{\int_a^b x^n U_{n-1}(x)\,d\dot{\psi}(x)}{\int_a^b x^{n-2}U_{n-2}(x)\,d\dot{\psi}(x)} = -\frac{1}{a_n^2}\,\frac{a_{n-1}a_{n-2}\phi_{n-2}(\alpha)}{\phi_{n-1}(\alpha)},$$

$$\bar{b}_{n+1} = \frac{\int_a^b x^2 U^2_{n-1}(x)\,d\dot{\psi}(x)}{\int_a^b U^2_{n-1}(x)\,d\dot{\psi}(x)} = \bar{s}_{n-1} - S_{n+1} = \bar{s}_{n-1} - \bar{s}_{n+1},$$

$$\bar{s}_{n+1} = S_{n+1}.$$

Multiplying (45) by $x^n(x-\alpha)\,d\psi(x)$, integrating, and making use of (10, 13, 2), we get further:

$$\left[S_{n+1}S_{n+2} - \frac{a_{n+2,n}}{a_{n+2}}\right] - \left[S_{n-1}S_n - \frac{a_{n,n-2}}{a_n}\right] + \frac{a_{n-2}}{a_{n-1}}\,\frac{\phi_{n-2}(\alpha)}{\phi_{n-1}(\alpha)} + (\bar{s}_{n-1} - S_{n+1})S_{n+1} + \bar{c}_{n+1} = 0.$$

Furthermore, using (12), we get

$$\bar{s}_{n-1} = S_{n-1} - \frac{a_{n-2}\phi_{n-2}(\alpha)}{a_{n-1}\phi_{n-1}(\alpha)},$$

$$\bar{c}_{n+1} = \left[S_{n-1}S_n - \frac{a_{n,n-2}}{a_n}\right] - \left[S_{n+1}S_{n+2} - \frac{a_{n+2,n}}{a_{n+2}}\right] + \frac{a_{n-2}\phi_{n-2}(\alpha)}{a_{n-1}\phi_{n-1}(\alpha)}\,c_{n+1} + S_{n+1}(S_{n+1} - S_{n-1}).$$

In the so-called "symmetric" case $(\psi(x) \equiv -\psi(-x);\ (a,b) \equiv (-h,h))$

$$S_n = c_n = 0 \quad (n \geqq 1), \quad \frac{a_{n,n-2}}{a_n} = -\sum_{i=2}^{n} \lambda_i,$$

so that

$$\bar{b}_{n+1} = \bar{s}_{n-1} = -\sqrt{\lambda_n}\,\frac{\phi_{n-2}(\alpha)}{\phi_{n-1}(\alpha)}, \qquad \bar{c}_{n+1} = -\lambda_{n+2} - \lambda_{n+1} \quad (\text{“symmetric” case}).$$

(iii) $\phi_n(\alpha) = \phi_{n-2}(\alpha) = 0$. Proceeding as before, we get:

$$l_{n+1} = \frac{a^2_{n-2}}{a_n^2}; \qquad \bar{b}_{n+1} = \bar{s}_{n-1} - \bar{s}_{n+1} = S_{n-1} - S_{n+1};$$

$$\bar{c}_{n+1} = \left[S_{n-1}S_n - \frac{a_{n,n-2}}{a_n}\right] - \Big[S_{n+1}S_{n+2} - \frac{a_{n+2,n}}{a_{n+2}} + S_{n+1}(S_{n+1} - S_{n-1});$$

$$\bar{b}_{n+1} = 0; \qquad \bar{c}_{n+1} = -\lambda_{n+2} - \lambda_{n+1} \quad (\text{“symmetric” case}).$$

If $\phi_n(\alpha) = 0$, $\phi_{n+2}(\alpha) \neq 0$, we write

$$U_{n+2}(x) = (x - k_{n+2}) U_{n+1}(x) + P_n(x), \text{ and } \int_a^b P_n(x) G_{n-1}(x)(x-\alpha) d\psi(x) = 0.$$

Hence,

$$(46) \quad P_n(x) = -l_{n+2} U_{n-1}(x), \quad U_{n+2}(x) = (x - k_{n+2}) U_{n+1}(x) - l_{n+2} U_{n-1}(x),$$

and as above

$$l_{n+2} = -\frac{a_n^2}{a^2_{n+2}} \frac{\phi_{n+2}(\alpha)}{\phi_{n+1}(\alpha)}, \qquad k_{n+2} = c_{n+2} - \frac{a_{n+1}}{a_{n+2}} \frac{\phi_{n+1}(\alpha)}{\phi_{n+2}(\alpha)}.$$

In order to illustrate we take in Theorem VI $r(x) \equiv x - \alpha$ and assume $\phi_1(\alpha)\phi_2(\alpha) \cdots \phi_{k-1}(\alpha) \neq 0$, $\phi_k(\alpha) = 0$. *Then the polynomials* $q_i(x)$ $(i = 1, 2, \cdots, k-1)$ *in* (41) *are all of the first degree, while* $q_k(x)$ *is of the second degree. Correspondingly, the recurrence relation* (19) *holds for* $(n = 1, 2, \cdots, k-1)$; *for* $n = k+1$ *its character changes as indicated under* (ii), (iii). (For $n = k$, (19) does not exist).

Case (iii) *is possible.* This is evident in the symmetric case with $\alpha = 0$, for here $\phi_{2n-1}(\alpha) = 0$ $(n = 1, 2, \cdots)$ so that all the $q_i(x)$ in (41) are of degree 2. This shows that the upper bound for the degree of $q_i(x)$ as given in Theorem VI is actually attained in this case.

14. *A minimum property of* $U_n(x)$. *Among all polynomials* $G_n(x)$ *such that* $G_n(\alpha) = 1$ $(\phi_n(\alpha) \neq 0)$, *it is the polynomial* $\dfrac{K_n(x, \alpha)}{K_n(\alpha)} = \dfrac{a_n\phi_n(\alpha)}{K_n(\alpha)} U_n(x)$ *which minimizes the integral* $\int_a^b G_n^2(x) d\psi(x)$, with the minimum $\dfrac{1}{K_n(\alpha; d\psi)}$. The proof can be easily accomplished by using the methods of constrained extrema.

15. *The case of two changes of sign.* We wish to investigate the existence of a system of polynomials

$$\{V_n(x) \equiv x^n + \cdots\} \qquad (n = 0, 1, \cdots)$$

satisfying the condition of orthogonality

$$(47) \quad \int_a^b V_n(x) G_{n-1}(x) d\bar{\psi}(x), \text{ with } \bar{\psi}(x) \equiv \int_a^x (t - \alpha_1)(t - \alpha_2) d\psi(t),$$
$$(a < \alpha_1 < \alpha_2 < b).$$

We get as above (see § 2), writing

$$(x - \alpha_1)(x - \alpha_2) V_n(x) \equiv \sum_{i=0}^{n+2} A_i \phi_i(x):$$

$$(48) \qquad \begin{cases} A_{n+1}\phi_{n+1}(\alpha_1) + A_n\phi_n(\alpha_1) = -\dfrac{\phi_{n+2}(\alpha_1)}{a_{n+2}} \\ A_{n+1}\phi_{n+1}(\alpha_2) + A_n\phi_n(\alpha_2) = -\dfrac{\phi_{n+2}(\alpha_2)}{a_{n+2}}. \end{cases}$$

The determinant of (48) is $\dfrac{a_{n+1}}{a_n}(\alpha_1 - \alpha_2)K_n(\alpha_1, \alpha_2)$.

The condition $K_n(\alpha_1, \alpha_2) \neq 0$ is thus seen to be sufficient for the existence of $V_n(x)$ satisfying (47). Furthermore, this condition insures the unique determination of $V_n(x)$ in the form

$$V_n(x) = \frac{a_n}{a_{n+1}a_{n+2}(\alpha_1 - \alpha_2)K_n(\alpha_1, \alpha_2)(x-\alpha_1)(x-\alpha_2)} \begin{vmatrix} \phi_{n+2}(x) & \phi_{n+1}(x) & \phi_n(x) \\ \phi_{n+2}(\alpha_1) & \phi_{n+1}(\alpha_1) & \phi_n(\alpha_1) \\ \phi_{n+2}(\alpha_2) & \phi_{n+1}(\alpha_2) & \phi_n(\alpha_2) \end{vmatrix}.$$

In particular, if $K_n(\alpha_1, \alpha_2) \neq 0$ and $\phi_{n+2}(\alpha_1) = \phi_{n+}(\alpha_2) = 0$,

$$V_n(x) = \frac{\phi_{n+2}(x)}{a_{n+2}(x-\alpha_1)(x-\alpha_2)}.$$

Moreover, in this latter case (47) is replaced by

$$\int_a^b V_n(x)G_{n+1}(x)\,d\psi(x) = 0.$$

On the other hand, if $K_n(\alpha_1, \alpha_2) = 0$, the consistency of the system (48) requires the matrix

$$\begin{Vmatrix} \phi_{n+2}(\alpha_1) & \phi_{n+1}(\alpha_1) & \phi_n(\alpha_1) \\ \phi_{n+2}(\alpha_2) & \phi_{n+1}(\alpha_2) & \phi_n(\alpha_2) \end{Vmatrix}$$

to be of rank one. We proceed to show that this is possible. In the first place, we must have

$$\begin{vmatrix} \phi_{n+2}(\alpha_1) & \phi_{n+1}(\alpha_1) \\ \phi_{n+2}(\alpha_2) & \phi_{n+1}(\alpha_2) \end{vmatrix} = \frac{a_{n+2}}{a_{n+1}}(\alpha_1 - \alpha_2)K_{n+1}(\alpha_1, \alpha_2) = 0, \quad K_{n+1}(\alpha_1, \alpha_2) = 0,$$

which, combined with the assumed relation $K_n(\alpha_1, \alpha_2) = 0$, gives

$$\phi_{n+1}(\alpha_1)\phi_{n+1}(\alpha_2) = 0.$$

If we assume now that $\phi_{n+1}(\alpha_1) = 0$ (hence $\phi_n(\alpha_1) \neq 0$), then (8) shows that $\phi_{n+1}(\alpha_2) = 0$. Conversely, assuming $\phi_{n+1}(\alpha_1) = \phi_{n+1}(\alpha_2) = 0$, we get at once $K_n(\alpha_1, \alpha_2) = K_{n+1}(\alpha_1, \alpha_2) = 0$.

In the second place, consider the last determinant of order 2 of our matrix and use (5, 2):

$$\begin{vmatrix} \phi_{n+2}(\alpha_1) & \phi_n(\alpha_1) \\ \phi_{n+2}(\alpha_2) & \phi_n(\alpha_2) \end{vmatrix} = \begin{vmatrix} (\alpha_1 - c_{n+2})\dfrac{a_{n+2}}{a_{n+1}}\phi_{n+1}(\alpha_1) & -\lambda_{n+2}\phi_n(\alpha_1) & \phi_n(\alpha_1) \\ (\alpha_2 - c_{n+2})\dfrac{a_{n+2}}{a_{n+1}}\phi_{n+1}(\alpha_2) & -\lambda_{n+2}\phi_n(\alpha_2) & \phi_n(\alpha_2) \end{vmatrix} = 0.$$

Hence, if $\phi_{n+1}(\alpha_1) = \phi_{n+1}(\alpha_2) = 0$, the determination of $V_n(x)$ by means of (48) is no longer unique. In this case A_{n+1} in (48) may be assigned arbitrarily. These considerations lead to

THEOREM VII. *A necessary and sufficient condition that a uniquely determined polynomial, of a given degree n,* $V_n(x) = x^n + \cdots$, *satisfying* (47), *exist, is:* $K_n(\alpha_1, \alpha_2) \neq 0$. *Moreover, if* $K_n(\alpha_1, \alpha_2) \neq 0$ *for all n,*[10] *there exists a uniquely determined sequence* $\{V_n(x)\}$ $(n = 0, 1, \cdots)$ *of such polynomials.*

The polynomial $V_n(x)$, may have two (but not more than two) imaginary or equal zeros, or two (but not more than two) zeros outside the interval (a, b). To show this construct the polynomial $\Phi_n(x; (x^2 + r^2)d\psi) = x^n + \cdots$, ($r$ arbitrary real constant), orthogonal with respect to the *monotonic* characteristic function $\int_a^x (t^2 + r^2)d\psi(t)$, and write the orthogonality property in the form

$$(49)\quad \int_a^b \frac{\Phi_n(x; (x^2 + r^2)d\psi)(x^2 + r^2)}{(x - \alpha_1)(x - \alpha_2)} G_{n-1}(x)(x - \alpha_1)(x - \alpha_2)d\psi(x) = 0,$$

where (α_1, α_2) are zeros of $\Phi_n(x; (x^2 + r^2)d\psi)$. (49) shows the existence of a polynomial

$$(50)\qquad V_n(x) \equiv \frac{\Phi_n(x; (x^2 + r^2)d\psi)(x^2 + r^2)}{(x - \alpha_1)(x - \alpha_2)} = x^n + \cdots,$$

orthogonal with respect to the characteristic function

$$\ddot{\psi}(x) = \int_a^x (t - a_1)(t - \alpha_2)d\psi(t).$$

Furthermore, we readily see that here the determination of $V_n(x)$ is unique since this is known to be true for $\Phi_n(x; (x^2 + r^2)d\psi)$. Hence, (50) shows that with such choice of α_1, α_2, $V_n(x)$ has two imaginary zeros. In like manner we show the possibility of the existence of two equal zeros or of two zeros outside (a, b).

The case of s $(s > 2)$ changes of sign could be treated as above. Since, however, even for two changes of sign the most important properties of the zeros of orthogonal polynomials corresponding to monotonic characteristic functions no longer hold, the discussion of this case is omitted.

UNIVERSITY OF PENNSYLVANIA.

[10] There is an infinity of such α_1, α_2 in any subinterval of (a, b).

METABELIAN GROUPS AND PENCILS OF BILINEAR FORMS.

By H. R. Brahana.

Introduction. In a recent paper [1] it was shown that the problem of classification of metabelian groups of order p^{n+m} which contain a given abelian group of order p^n as a maximal invariant abelian subgroup and have commutator subgroups of order p^m is equivalent to the problem of classification of the matrices $x_1M_1 + x_2M_2 + \cdots + x_kM_k$ under projective transformations on the x's and elementary transformations on the square matrices $M_1, M_2, \cdots, M_k$. The x's and the elements of the M's are of course numbers in a modular field as are also the coefficients of the transformations. The squareness of the matrices comes from the requirement that the commutator subgroup be of order p^m. The situation may then be discussed in terms of the invariant factors of the matrix $x_1M_1 + x_2M_2 + \cdots + x_kM_k$.

The argument of that paper still holds when the commutator subgroup is not of order p^m and the matrices M_i are not square. In this case however we are deprived of the use of a well-developed theory of invariant factors. So far as I know the question of the conjugacy of two matrices of the above type under transformation on the x's and simultaneous transformations on rectangular M's has not been considered. It is our purpose to consider the groups which give rise to such matrices in the simple case where $m = 4$ and $k = 2$ and to use the results to obtain normal forms for the matrices. It will be convenient to interpret the matrices M_i as matrices of bilinear forms in which case the matrix above, which we shall denote hereafter as $\lambda_1M_1 + \lambda_2M_2$, may be taken to represent a pencil of bilinear forms.

1. *The groups.* We consider groups $G = \{H, U\}$ where H is abelian, of order p^n, and type $1, 1, \cdots$ and U is an abelian group of order p^4 and type $1, 1, \cdots$ from the group of isomorphisms of H. We require that no operator of U, except identity, be permutable with every operator of H. We require further that G be metabelian which implies that its commutator subgroup is in its central, that every operator of U determines [2] a partition of n with greatest term equal to 2. Finally, we require that no operator of U

[1] "Metabelian groups of order p^{n+m} with commutator subgroups of order p^m," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 776-792.

[2] "On metabelian groups," *American Journal of Mathematics*, vol. 56 (1934), pp. 490-510.

determine a partition of n in which more than two terms are equal to 2. An operator U_i will be said to be of type I or type II according as the partition it determines contains one or two 2's. The group $\{H, U_i\}$ will be said to be of type I or type II depending on the type of U_i. In the groups which we shall consider U will contain only operators of types I and II, the identity excepted. Since the groups in which U contains only operators of type I were classified in the paper just referred to we shall suppose that U contains at least one operator of type II.

The central of G under these conditions is of order p^{n-2} and we may suppose that generators of H are chosen so that all but two are in the central. Let the two of the generators of H which are not in the central be denoted by s_1 and s_2. Then if U_i is an operator of U, $\{H, U_i\}$ will have a commutator subgroup of order p or p^2 according as it is of type I or type II. The maximum order for the commutator subgroup of G is p^8 and occurs only if each of the operators of every set of four which generate U is of type II and the resulting eight commutators are independent. The commutator subgroup of G has an order at least p^2 since U contains at least one operator of type II. Since the commutator subgroup is characteristic we may separate the groups in question into classes according to the orders of their commutator subgroups and no two groups belonging to different classes can be simply isomorphic.

Let the order of the commutator subgroup be p^l. It is immediately obvious that there is but one group in the class corresponding to $l = 8$, for a simple isomorphism is established between any two such groups by a proper naming of generators. These considerations give the following more general theorem:

(1.1) *There is but one metabelian group* $G = \{H, U\}$ *of order* p^{n+m} *with commutator subgroup of order* p^{2m}, *provided the operators of* U *are restricted to types I and II. In this group no operator of* U *is of type I.*

If two groups U and U' have different numbers of subgroups composed of operators of type I, there will exist no simple isomorphism between $\{H, U\}$ and $\{H, U'\}$ in which H corresponds to itself.[3] Accordingly, for $l < 8$ we may consider the groups in sets determined by the number of subgroups of type I in U.

When $l = 7$, U cannot contain more than one subgroup of type I. Otherwise two operators U_1 and U_2 both of type I could be selected as two of the

[3] We shall postpone the question of simple isomorphisms between $\{H, U\}$ and $\{H, U'\}$ in which H does not correspond to itself.

four independent generators of U. The group $\{H, U_1, U_2\}$ would have a commutator subgroup of order p^2. The commutator subgroup of $\{H, U_3, U_4\}$ could be of order at most p^4 and hence the commutator subgroup of G could be of order at most p^6. If U contains one subgroup of type I, let us suppose it to be generated by U_1 and generators of H chosen so that U_1 is permutable with s_2. Then $\{H, U_2, U_3, U_4\}$ must have a commutator subgroup of order p^6 and by (1. 1) just one such group exists. Consequently there exists such a group and it is completely determined by the requirements that $l = 7$ and that U contain but one subgroup of type I.[4] If U contains no operator of type I there is also but one group. No matter how $U_1, \cdots, U_4$ are chosen, so long as they generate U, the group $\{H, U_1, U_2\}$ will have a commutator subgroup of order p^3 or p^4 and in the former case $\{H, U_3, U_4\}$ will have a commutator subgroup of order p^4. We may then assume that the commutator subgroup K' of $\{H, U_1, U_2\}$ is of order p^4. Then at most one of the commutators arising from transformation of H by U_3 is in K' and if so the two commutators arising from U_4 are independent of K'. Hence we may choose an operator U_3 such that the commutator subgroup K'' of $\{H, U_1, U_2, U_3\}$ is of order p^6. Then of the two commutators arising from transformation of H by U_4 at most one is in K''. From the symmetry in s_1 and s_2 of the relations which generators of G just described satisfy it is clear that no restriction is introduced by assuming that the commutator of U_4 and s_1 is in K''. Let us denote this commutator by s_k, and the commutator of U_4 and s_2 by s_9. If s_k is not in K'' it is in $\{K'', s_9\}$. If then we replace s_1 by a proper combination of s_1 and s_2 we obtain a commutator s'_k which is in K''. We may assume further that s_k is in the part of K'' which arises from transformation of s_2 by U_1, U_2, and U_3, for otherwise U_4 could be replaced by such a combination of U_1, U_2, U_3, and U_4 that such would be the case. Therefore there exists in $\{U_1, U_2, U_3\}$ an operator U' whose commutator with s_2 is s_k. This operator may be taken to be U_1, and consequently we may assume that the commutator of U_1 and s_2 is the same as that of U_4 and s_1. There are therefore two groups for $l = 7$ and they are distinguished by the numbers of subgroups of type I in U. These considerations also apply more generally to give the theorem:

(1. 2) *If the operators of U are all of types I and II, there are two and only two groups of the type we are considering of order p^{n+m} with commutator subgroups of order p^{2m-1}. In one of them U contains one subgroup of type I, and in the other none.*

---

[4] We use the shorter expression "subgroup of type I" in place of "subgroup composed of operators of type I, and the identity."

When $l=6$ the number of independent subgroups of type I in U can be at most two as may be seen from an argument similar to that used for $l=7$. We shall see that U may contain more than two subgroups of type I but that if so all such subgroups are contained in the group generated by two of them. The possibilities for the number of independent subgroups of type I are therefore 2, 1, and 0.

If U contains two independent subgroups of type I, let them be generated by U_3 and U_4. Two possibilities arise: the subgroups of $\{s_1, s_2\}$ permutable with the respective operators U_3 and U_4 may be the same or they may be different. If they are the same every operator of $\{U_3, U_4\}$ is of type I; if they are different every operator of $\{U_3, U_4\}$ except those in $\{U_3\}$ and $\{U_4\}$ is of type II. In either case since the commutator subgroup of $\{H, U_1, U_2\}$ is of order at most p^4, the commutator subgroup of $\{H, U_3, U_4\}$ must be of order p^2. In either case every operator of U not in $\{U_3, U_4\}$ is of type II. Hence, in the one case U contains $1+p$ subgroups of type I, and in the other contains two subgroups of type I. The two groups are generated by operators $s_1, s_2, s_3, \cdots, s_n, U_1, \cdots, U_4$ which satisfy the following relations: when U contains $1+p$ subgroups of type I,

$$(1.3)\quad \begin{aligned} &U_1^{-1}s_1U_1=s_1s_3,\ U_2^{-1}s_1U_2=s_1s_5,\ U_3^{-1}s_1U_3=s_1s_7,\ U_4^{-1}s_1U_4=s_1s_8,\\ &U_1^{-1}s_2U_1=s_2s_4,\ U_2^{-1}s_2U_2=s_2s_6; \end{aligned}$$

when U contains two subgroups of type I,

$$\begin{aligned} &U_1^{-1}s_1U_1=s_1s_3,\ U_2^{-1}s_1U_2=s_1s_5,\ U_3^{-1}s_1U_3=s_1s_7,\\ &U_1^{-1}s_2U_1=s_2s_4,\ U_3^{-1}s_2U_3=s_2s_6,\qquad U_4^{-1}s_2U_4=s_2s_8. \end{aligned}$$

If U contains one subgroup of type I let it be generated by U_4 and let U_4 be permutable with s_2. The commutator subgroup of $\{H, U_1, U_2, U_3\}$ must be of order p^5 or p^6, and $\{U_1, U_2, U_3\}$ can contain no operator of type I. By theorems (1.1) and (1.2) there is but one group in each case. In the former case the group is completely determined, since the commutators of $\{H, U_4\}$ and $\{H, U_1, U_2, U_3\}$ are independent. A set of generating relations is obtained by adding $U_3^{-1}s_2U_3=s_2s_3$ to (1.3) above. In the second case we may obtain such a group by adding the relation $U_3^{-1}s_2U_3=s_2s_k$ to (1.3). It then follows that s_8 is in the group $\{s_3, s_4, s_5, s_6, s_7, s_k\}$. The group $\{s_3, s_5, s_7, s_8\}$ must be of order p^4, otherwise a proper combination U' of U_1, U_2, and U_3 would transform s_1 into s_1s_8 and $U'U_4^{-1}$ would be of type I contrary to the assumption that U contained no operator of type I besides U_4. At least two of the operators s_4, s_6, and s_k are independent of $\{s_3, s_5, s_7, s_8\}$; we may then suppose that s_8 is expressible in terms of s_3, s_4, s_5, s_6, s_7, and s_k.

If the group which we are considering is distinct from the one just previously described, we must expect at least two possibilities to appear, for if in the former group we replace U_1 by U_1U_4 the commutator subgroup of $\{H, U_1, U_2, U_3\}$ is of order p^6, being generated by $s_3s_8, s_4, s_5, s_6, s_7$, and s_8. There are then at least these two possibilities in the present case: (a) s_8 is in the group $\{s_3, s_4, s_5, s_6, s_7, s_k\}$ and not in $\{s_4, s_6, s_k\}$, or (b) s_8 is in $\{s_4, s_6, s_k\}$. In case (b) we may assume that $s_k = s_8$, for every operator of $\{s_4, s_6, s_k\}$ is a commutator, and any set of three independent operators of $\{U_1, U_2, U_3\}$ generate it. The operator U_4 is a unique operator in U, it defines s_2 to within a power of a single operator of $\{s_1, s_2\}$, and s_2 in turn defines $\{s_4, s_6, s_k\}$. Hence, this group is distinct from the one previously defined, in which s_8 is not in $\{s_4, s_6, s_k\}$ and as we have seen may lead to case (a). We complete this by showing that case (a) leads to a single group, consequently one in which U_1, U_2, U_3 can be chosen so that $\{H, U_1, U_2, U_3\}$ has a commutator subgroup of order p^5. If s_8 is in $\{s_3, \cdots, s_7, s_k\}$ but not in $\{s_4, s_6, s_k\}$, then there exists in $\{U_1, U_2, U_3\}$ an operator U'_1 and in $\{s_1, s_2\}$ an operator s'_1 such that the commutator of U'_1 and s'_1 is s_ks_8. This operator s'_1 is obviously not a power of s_2. The operator U'_1 is not U_3 for in that case $U_3U_4^{-1}$ would give the same commutator with s'_1 as with s_2 and hence would be of type I. Consequently, if we replace U_1 by $U'_1U_4^{-1}$ and s_1 by s'_1, generators of the group satisfy the relations found for the first group with one subgroup of type I.

The two groups each containing one subgroup of type I just described have been distinguished by the order of the group $\{s_4, s_6, s_k, s_8\}$ which in one case is p^3 and in the other p^4. They may also be distinguished from each other by the non-abelian subgroups which they contain. In one there exists at least one subgroup $\{H, U_1, U_2, U_3\}$ of order p^{n+3} with commutator subgroup of order p^5 which contains no subgroup of type I; in the other every subgroup of order p^{n+3} with commutator subgroup of order p^5 contains a subgroup of type I. Thus there is no simple isomorphism between the two groups in which H corresponds to H.[5]

Now suppose that U contains no operator of type I. Two possibilities are immediately evident: (a) U contains two operators U_1 and U_2 such that $\{H, U_1, U_2\}$ has a commutator subgroup of order p^2; or (b) U contains no

[5] We beg leave to point out that all the operators of type II in the group of isomorphisms of H are conjugate, as are all the operators of type I. The two groups U and U' are abelian, of order p^4, and type $1, 1, \cdots$ and every operator of one can be transformed into many operators of the other by operators which transform H into itself. U may be transformed into U' in many ways. U may not, however, be transformed into U' by any operator which leaves H invariant.

such group. The condition (a) completely determines the group, for the commutator subgroup of $\{H, U_3, U_4\}$ must then be of order p^4 and must be independent of the commutator subgroup of $\{H, U_1, U_2\}$. In any case, no matter how s_1 and s_2 are chosen the commutator subgroups H_1 and H_2 arising from transformation of s_1 and s_2 respectively by U are of order p^4, since U contains no operator of type I. Since $\{H_1, H_2\}$ is of order p^6, these two groups have a subgroup of order p^2 in common. This subgroup determines two operators U_1 and U_2 such that the commutator subgroup arising from transformation of s_1 by $\{U_1, U_2\}$ is the subgroup common to H_1 and H_2; it determines two other operators V_1 and V_2 such that the commutator subgroup arising from transformation of s_2 by $\{V_1, V_2\}$ is the same group. In case (a) the group $\{U_1, U_2, V_1, V_2\}$ is of order p^2. In case (b) this group is of order at least p^3. In case it is of order p^4, the four commutators arising from transformation of s_1 by V_1 and V_2 and of s_2 by U_1 and U_2 must be independent of the commutator subgroup common to H_1 and H_2. The group is therefore completely determined. If the group $\{U_1, U_2, V_1, V_2\}$ is of order p^3, it is generated by three operators U_1, U_2, and U_3 and its commutator subgroup is of order p^4. Then the commutator subgroup of $\{H, U_4\}$ must be independent of this group of order p^4, and since there is but one group of order p^{n+3} with commutator subgroup of order p^4 and no subgroups of type I [6] it is completely determined. It is necessary to determine whether or not these last two groups are distinct. This last group contains a subgroup of order p^{n+3} with commutator subgroup of order p^4 and the preceding group does not.

Normal forms for generating relations of the three groups with commutator subgroups of order p^6 and no subgroups of type I are:

$$U_1^{-1}s_1U_1 = s_1s_3,\quad U_2^{-1}s_1U_2 = s_1s_4^r,\quad U_3^{-1}s_1U_3 = s_1s_5,\quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4,\quad U_2^{-1}s_2U_2 = s_2s_3,\quad U_3^{-1}s_2U_3 = s_2s_6,\quad U_4^{-1}s_2U_4 = s_2s_8,$$

where r is any not-square. This group contains a subgroup of order p^{n+2} with commutator subgroup of order p^2.

$$U_1^{-1}s_1U_1 = s_1s_3,\quad U_2^{-1}s_1U_2 = s_1s_5,\quad U_3^{-1}s_1U_3 = s_1s_6,\quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_2^{-1}s_2U_1 = s_2s_4,\quad U_2^{-1}s_2U_2 = s_2s_3,\quad U_3^{-1}s_2U_3 = s_2s_5,\quad U_4^{-1}s_2U_4 = s_2s_8.$$

This group contains no subgroup of order p^{n+2} with commutator subgroup of order p^2, but contains one of order p^{n+3} with commutator subgroup of order p^4.

$$U_1^{-1}s_1U_1 = s_1s_3,\quad U_2^{-1}s_1U_2 = s_1s_5,\quad U_3^{-1}s_1U_3 = s_1s_7,\quad U_4^{-1}s_1U_4 = s_1s_8,$$
$$U_1^{-1}s_2U_1 = s_2s_4,\quad U_2^{-1}s_2U_2 = s_2s_6,\quad U_3^{-1}s_2U_3 = s_2s_3,\quad U_4^{-1}s_2U_4 = s_2s_5.$$

[6] "On metabelian groups," *loc. cit.*, p. 510.

This group contains no subgroup of order p^{n+3} with commutator subgroup of order p^4.

As in the cases of $l = 8$ and $l = 7$, an obvious rewording of the argument above gives the more general theorem:

(1.4) *If the operators of U are all of types I and II there are seven and only seven groups of the type we are considering of order p^{n+m} with commutator subgroups of order p^{2m-2}. They may be characterized by their non-abelian subgroups of orders p^{n+1}, p^{n+2}, and p^{n+3}.*

When $l = 5$, the number of independent subgroups of type I in U cannot be greater than 3, and may be 3, 2, 1, or 0. If there are 3 independent subgroups of type I in U we may take them to be generated by U_2, U_3, and U_4. The operator U_1 must be of type II. The commutator subgroup of $\{H, U_2, U_3, U_4\}$ must be of order p^3 and must have no operator in common with the commutator subgroup of $\{H, U_1\}$. The groups will then be characterized by the properties of $\{H, U_2, U_3, U_4\}$. Each of the operators U_2, U_3, and U_4 is permutable with a subgroup of order p of $\{s_1, s_2\}$. These three subgroups may be the same; two of them may coincide; or all three may be distinct. If the three subgroups coincide let us suppose the subgroup to be generated by s_2. Then generators of G will satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_6, \quad U_4^{-1}s_1U_4 = s_1s_7.$$
$$U_1^{-1}s_2U_1 = s_2s_4.$$

In this case every operator of $\{U_2, U_3, U_4\}$ will be of type I, and G will contain $1 + p + p^2$ subgroups of type I.

In the second case the two subgroups of $\{s_1, s_2\}$ permutable with operators of type I of U may be taken to be generated by s_1 and s_2. In this case generators of G will satisfy relations the same as those above except that $U_4^{-1}s_1U_4 = s_1s_7$ is replaced by $U_4^{-1}s_2U_4 = s_2s_7$. The only operators of type I in $\{U_2, U_3, U_4\}$ are the operators of $\{U_2, U_3\}$ and powers of U_4. G therefore contains $2 + p$ subgroups of type I.

In the third case we may suppose that two of the groups are generated by s_1 and s_2 respectively. Generators of G will satisfy the relations:

$$\begin{aligned} &U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_6, \\ &U_1^{-1}s_2U_1 = s_2s_4, \qquad\qquad\qquad\qquad\quad U_3^{-1}s_2U_3 = s_2s_6, \quad U_4^{-1}s_2U_4 = s_2s_7. \end{aligned}$$

It is obvious that the group $\{U_2, U_3, U_4\}$ contains but three subgroups of type I, and that G likewise contains but three subgroups of type I.

When U contains but two independent subgroups of type I the operators of $\{s_1, s_2\}$ permutable respectively with these two subgroups of U may constitute one or two subgroups of order p. If they constitute one subgroup of order p, let it be generated by s_2. Let U_3 and U_4 generate the respective subgroups of type I. Then $\{U_1, U_2\}$ can contain only operators of type II. The commutator subgroup of $\{H, U_1, U_2\}$ is of order p^3 or p^4. There is but one group for each of the orders p^3 and p^4 having the required properties. If the order of the commutator subgroup of $\{H, U_1, U_2\}$ is p^3 the group G is completely determined. Its generators satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_6, \quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3.$$

If the commutator subgroup of $\{H, U_1, U_2\}$ is of order p^4, then it must have a subgroup of order p in common with the commutator subgroup of $\{H, U_3, U_4\}$. This common subgroup may be taken to be in the group of commutators arising from transformation of s_2 by U_1 and U_2, for otherwise U'_1 and U'_2 could be chosen so that the commutator subgroup of $\{H, U'_1, U'_2\}$ would be of order p^3. The group G in this case is also completely determined. Its generators satisfy relations obtained from these above by changing the commutator of U_2 and s_2 from s_3 to s_6. The two groups are obviously distinct; each contains $1 + p$ subgroups of type I. They may be distinguished by the fact that the first contains a subgroup of order p^{n+2} with commutator subgroup of order p^3 and no subgroup of type I, whereas in the second every subgroup of order p^{n+2} with commutator subgroup of order p^3 contains subgroups of type I.

If the subgroups of $\{s_1, s_2\}$ permutable respectively with U_3 and U_4 are distinct let them be generated by s_1 and s_2. The commutator subgroups of $\{H, U_3\}$ and $\{H, U_4\}$ may coincide, in which case the commutator subgroup of $\{H, U_1, U_2\}$ must be of order p^4 and independent of it. Generators of G therefore satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_6, \qquad U_4^{-1}s_2U_4 = s_2s_7.$$

In this case G contains $1 + p$ subgroups of type I. The group is obviously distinct from the three preceding ones.

If the commutator subgroup of $\{H, U_3, U_4\}$ is of order p^2, there are but two subgroups of type I in $\{U_3, U_4\}$ and therefore but two in U. The commutator subgroup of $\{H, U_1, U_2\}$ is of order p^3 or p^4. In the first case generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_6,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3, \quad U_4^{-1}s_2U_4 = s_2s_7.$$

In the other case generators of G satisfy the above relations with s_3 replaced by s_k in the transform of s_2 by U_2. The operator s_k cannot be in the group $\{s_3, s_4, s_5\}$, but must be in the group $\{s_3, s_4, s_5, s_6, s_7\}$. We may assume that the expression for s_k in terms of these operators does not contain s_4 or s_7. It is therefore in the group $\{s_3, s_5, s_6\}$. There is an operator in the group $\{U_1, U_2, U_3\}$ whose commutator with s_1 is s_k, and this operator is not U_2. If it is not U_3, it may serve in place of U_1 and generators of G satisfy the relations above. If it is U_3, we have a new group whose generators satisfy the above relations with s_6 for the commutator of U_2 and s_2. This last group contains no subgroup of order p^{n+2} with commutator subgroup of order p^3 except those which contain subgroups of type I; the former group does contain such subgroups.

When U contains but one subgroup of type I let it be generated by U_4 and let the subgroup of $\{s_1, s_2\}$ permutable with U_4 be generated by s_2. Let $U_4^{-1}s_1U_4 = s_1s_7$. The commutator subgroup of $\{H, U_1, U_2, U_3\}$ is then of order p^4 or p^5. If it is of order p^4 it does not contain s_7. The two groups of commutators, H_1 and H_2, obtained by transforming s_1 and s_2 respectively by $\{U_1, U_2, U_3\}$ have a subgroup of order p^2 in common. This subgroup may be taken to be $\{s_3, s_4\}$ and it defines two operators U_1 and U_2 which give commutators s_3 and s_4 with s_1 and two operators V_1 and V_2 which give commutators s_3 and s_4 with s_2. The group $\{U_1, U_2, V_1, V_2\}$ is of order p^2 or p^3. We have the following two groups:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4, \quad U_3^{-1}s_1U_3 = s_1s_5, \quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3^r, \quad U_3^{-1}s_2U_3 = s_2s_6.$$

This group contains a subgroup of order p^{n+2} with commutator subgroup of order p^2.

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4, \quad U_3^{-1}s_1U_3 = s_1s_6, \quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_5, \quad U_2^{-1}s_2U_2 = s_2s_3, \quad U_3^{-1}s_2U_3 = s_2s_4.$$

This group contains no subgroup of order p^{n+2} with commutator subgroup of order p^2. These are the only two groups when U_1, U_2, and U_3 can be selected so that s_7 is not in the commutator subgroup of $\{H, U_1, U_2, U_3\}$.

If G contains no subgroup $\{H, U_1, U_2, U_3\}$ with commutator subgroup of order p^4, we may assume that H_1 and H_2 have a subgroup of order p in common. Generators of G will satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_6, \quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3, \quad U_3^{-1}s_2U_3 = s_2s_k.$$

The operator s_k is in the group $\{s_3, \cdots, s_7\}$ and it is not in the group $\{s_3, s_4, s_5, s_6\}$. In the expression for s_k neither s_3 nor s_4 need appear. It may therefore be assumed to be in the group $\{s_5, s_6, s_7\}$, but not in the group $\{s_5, s_6\}$. There exists in $\{U_2, U_3, U_4\}$ an operator U'_2 whose commutator with s_1 is s_k. This operator is not U_3 and may be used in place of U_2 if it is not U_4. If it is not U_4 then H_1 and H_2 have the group $\{s_3, s_k\}$ of order p^2 in common, which brings us back to the group previously described. We may then assume in this case that $s_k = s_7$. There are thus three distinct groups with commutator subgroup of order p^5, each containing one subgroup of type I. They are distinguished by means of their subgroups of orders p^{n+2} and p^{n+3}.

When U contains no operator of type I the two groups H_1 and H_2 are each of order p^4 and consequently have a subgroup of order p^3 in common. This subgroup determines three operators U_1, U_2, and U_3 such that the commutator subgroup arising from transformation of s_1 by $\{U_1, U_2, U_3\}$ is the common subgroup of H_1 and H_2. There are likewise three operators V_1, V_2, and V_3 determined by s_2 and the common subgroup. The order of $\{U_1, \cdots, V_3\}$ is either p^3 or p^4. If it is of order p^3 it is generated by U_1, U_2, and U_3. There is but one [7] such group $\{H, U_1, U_2, U_3\}$ with commutator subgroup of order p^3, and since its commutator subgroup must be independent of that of $\{H, U_4\}$ the group G is completely determined. Its generators satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_3^{\alpha}s_4^{\beta}, \quad U_4^{-1}s_1U_4 = s_1s_6$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3, \quad U_3^{-1}s_2U_3 = s_2s_5, \quad U_4^{-1}s_2U_4 = s_2s_7,$$

where $x^3 - \alpha x + \beta \equiv 0$ is irreducible, mod p.

If the order of $\{U_1, \cdots, V_3\}$ is p^4 two possibilities arise according as $\{U_1, \cdots, V_3\}$ does or does not contain a subgroup such that $\{H, U'_1, U'_2\}$ has a commutator subgroup of order p^2. In the first case generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4^{r}, \quad U_3^{-1}s_1U_3 = s_1s_5, \quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3, \quad U_3^{-1}s_2U_3 = s_2s_6, \quad U_4^{-1}s_2U_4 = s_2s_5.$$

In the second case generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_6, \quad U_4^{-1}s_1U_4 = s_1s_7,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3, \quad U_3^{-1}s_2U_3 = s_2s_5, \quad U_4^{-1}s_2U_4 = s_2s_6.$$

There are thus thirteen groups for $l = 5$.

[7] Cf. "On metabelian groups," *loc. cit.*

In the case where $l = 4$ we shall use the results of the paper [8] on metabelian groups of order p^{n+m} with commutator subgroups of order p^m. Two matrices M and N are used to describe the commutator structure of $\{s_1, s_3, s_4, \cdots, s_n, U_1, \cdots, U_4\}$ and $\{s_2, s_3, \cdots, s_n, U_1, \cdots, U_4\}$ respectively. If the commutator subgroup of G is $\{s_3, s_4, s_5, s_6\}$, the element m_{ij} of M is the exponent of s_{j+2} in the commutator of U_i and s_1; and the element n_{ij} of N is the exponent of s_{j+2} in the commutator of U_i and s_2. The groups G may then be classified according to the classification of the matrices $M + \lambda N$ under elementary transformations on the matrices M and N simultaneously and projective transformations on λ, both sorts of transformation having coefficients in the modular field, mod p.

Let us consider the determinant $f(\lambda) = | M + \lambda N |$, and suppose that $f(\lambda)$ is the product of four linear factors. If λ_1 is a root of $f(\lambda) \equiv 0$ then $s_1 s_2^{\lambda_1}$ is permutable with some operator of U. If all four roots of $f(\lambda) \equiv 0$ were the same we could take λ_1 to be zero and every element of M would be zero. This would imply that s_1 was permutable with every operator of U in which case U would contain only operators of type I. Since we assume U to contain at least one operator of type II, $f(\lambda)$ cannot be the fourth power of a linear expression in λ. We may then suppose that at least two of the linear factors of $f(\lambda)$ are distinct and that their zeros are 0 and ∞. The determinants of M and N are then both zero. If we consider first the case where $f(\lambda)$ has just two distinct zeros we have the two possibilities (a) $f(\lambda) = \lambda^3$ and (b) $f(\lambda) = \lambda^2$. In the first case three of the U's are permutable with s_1 and one with s_2 so that G contains $2 + p + p^2$ subgroups of type I. In the second case two U's are permutable with each so that G contains $2(1 + p)$ subgroups of type I.

When $f(\lambda)$ is the product of four linear factors three of which are distinct we may suppose its zeros to be 0, 1, and ∞ with 0 counted twice. Then two U's are permutable with s_1. In this case G contains $3 + p$ subgroups of type I.

When $f(\lambda)$ has four distinct zeros they may be taken to be 0, 1, ∞, and ρ, where ρ is the cross-ratio of the four taken in some arbitrary order. There are as many such groups as there are projectively distinct unordered sets of four points on the finite line, mod p. The cases where $\rho = 0$, 1, and ∞ give the group described just above. The values 2, -1, $1/2$ give a single group, and the two primitive cube roots of -1, when they exist in the modular field, give a single group. The rest of the numbers in the modular field go in sets of six to determine a single group. There are therefore $(p + 1)/6$ or

[8] *Loc. cit.*

13

$(p+5)/6$, according as p is of the form $6k-1$ or $6k+1$, groups which are distinct from each other and from the groups considered above. Each has four subgroups of type I.

When $f(\lambda)$ does not have four linear factors it has at most two. Let us consider the case where it has just two linear factors. If they are the same, we may suppose that $f(\lambda) = q(\lambda)\cdot\lambda^2$, where $q(\lambda)$ is an irreducible quadratic. All such quartics are conjugate under the projective group on λ, and hence there is but one such group. The group G contains $1+p$ subgroups of type I, and it contains a subgroup of order p^{n+2} with commutator subgroup of order p^2 and no operators of type I; this last subgroup corresponds to the irreducible quadratic.

If the two linear factors of $f(\lambda)$ are distinct we have $f(\lambda) = q(\lambda)\cdot\lambda$. U contains one subgroup permutable with s_1 and one permutable with s_2. G contains therefore two subgroups of type I and a subgroup of order p^{n+2} with commutator subgroup of order p^2 and no operator of type I. The number of distinct such groups is the number of conjugate sets of polynomials $q(\lambda)(\lambda-\lambda_1)(\lambda-\lambda_2)$ under the "rational" projective group on λ. The irreducible quadratic may be transformed into any particular quadratic and there is then a group of order $2(p+1)$ which leaves it fixed. There exists an operator of order two which leaves $q(\lambda)$ and $f(\lambda)$ fixed, viz., the one which interchanges the zeros of $q(\lambda)$ and also interchanges λ_1 and λ_2. Consequently there are $p+1$ quadratics $(\lambda-\lambda_1)(\lambda-\lambda_2)$ such that $q(\lambda)(\lambda-\lambda_1)(\lambda-\lambda_2)$ belong to the same conjugate set unless λ_1, λ_2, and the zeros of $q(\lambda)$ constitute a harmonic set in which case there are $(p+1)/2$ in the conjugate set. Of the first kind there are therefore $(p-1)/2$ conjugate sets and of the second kind one. There are in all $(p+1)/2$ groups of this kind.

When $f(\lambda)$ has but one linear factor the other factor is an irreducible cubic. G has one subgroup of type one, and a subgroup of order p^{n+3} with commutator subgroup of order p^3 and no subgroup of type I. To determine the number of such groups, let us write $f(\lambda) = c(\lambda)(\lambda-\lambda_1)$, where $c(\lambda)$ is the irreducible cubic. The "rational" projective group contains an operator of order three which transforms a root λ_0 of $c(\lambda)\equiv 0$ into its p-th power.[9] This group associates three numbers λ_1, λ_2, and λ_3 with any one of them, so that the cross-ratio of $\lambda_0, \lambda_0^p, \lambda_0^{p^2}, \lambda_1$ equals the cross-ratio of $\lambda_0^p, \lambda_0^{p^2}, \lambda_0, \lambda_2$ equals the cross-ratio of $\lambda_0^{p^2}, \lambda_0, \lambda_0^p, \lambda_3$. There are therefore $(p+5)/3$ or $(p+1)/3$ conjugate sets of such quartics depending on whether p is of the form $6k+1$ or $6k-1$.[10]

---

[9] Cf. "On cubic congruences," *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 962-969.

[10] The cubic $c(\lambda)$ appears as a quartic with one root infinite. And when

There remain the groups G with no subgroups of type I. There are $p+2$ such groups. There are thus $2p+7$ or $2p+9$ groups with commutator subgroups of order p^4 according as p is of the form $6k-1$ or $6k+1$.

Let us now suppose that $l=3$. Then however generators of $\{s_1, s_2\}$ are selected, H_1 and H_2 are of orders not greater than p^3. Consequently, U contains at least two subgroups of type I. Let us suppose that U contains four independent operators of type I. The operators of $\{s_1, s_2\}$ permutable with them must constitute at least two subgroups of order p. If they constitute two subgroups, then two possibilities arise: (a) three of the four independent U's are permutable with s_1 and the other with s_2, or (b) two U's are permutable with s_1 and two with s_2. In case (a) generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4, \quad U_3^{-1}s_1U_3 = s_1s_5,$$
$$U_4^{-1}s_2U_4 = s_2s_5.$$

In this case the group $\{U_1, U_2, U_3\}$ contains only operators of type I, as does $\{U_3, U_4\}$. There are therefore $1+p+p^2$ such subgroups in the first, and $1+p$ in the second; only the subgroup generated by U_3 is in both. Hence G contains $1+2p+p^2$ subgroups of type I. In case (b) generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4,$$
$$U_3^{-1}s_2U_3 = s_2s_4, \quad U_4^{-1}s_2U_4 = s_2s_5.$$

In this case $\{U_1, U_2\}$, $\{U_2, U_3\}$, and $\{U_3, U_4\}$ each contain $1+p$ subgroups of type I of which two, generated by U_2 and U_3 respectively, are counted twice each. G therefore contains $1+3p$ subgroups of type I.

When the four independent U's of type I are permutable with three subgroups of order p of $\{s_1, s_2\}$, we may suppose two of them to be permutable with s_1, one with s_2, and the other with $s_1s_2^{-1}$. The group H_1 must be of order p^3, otherwise U_3 could be selected so that it was permutable with s_1. Generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4, \quad U_3^{-1}s_1U_3 = s_1s_5,$$
$$U_3^{-1}s_2U_3 = s_2s_5, \quad U_4^{-1}s_2U_4 = s_2s_3.$$

G contains $1+3p$ subgroups of type I.

$p=6k+1$, $c(\lambda)$ may be taken to be λ^3+a, where $-a$ is not a cube; the group leaving $c(\lambda)$ fixed is generated by $\lambda'=\rho\lambda$, where ρ is a primitive cube root of unity. This leaves $c(\lambda)$ and $c(\lambda)\cdot\lambda$ fixed; the remaining $6k$ quartics are separated into $2k$ sets of 3 each. We have thus the $(p+5)/3$ sets.

When the four independent U's are permutable with different subgroups of $\{s_1, s_2\}$ we may suppose U_1 permutable with s_2, U_2 with $s_1s_2^{-1}$, U_3 with $s_1s_2^{-r}$, and U_4 with s_1. Generators of G satisfy the relations:

$$
\begin{array}{llll}
U_1^{-1}s_1U_1 = s_1s_3, & U_2^{-1}s_1U_2 = s_1s_4, & U_3^{-1}s_1U_3 = s_1s_5^r, & \\
 & U_2^{-1}s_2U_2 = s_2s_4, & U_3^{-1}s_2U_3 = s_2s_5, & U_4^{-1}s_2U_4 = s_2s_3.
\end{array}
$$

The groups $\{U_1, U_4\}$, $\{U_1U_4, U_2\}$, $\{U_1^rU_4, U_3\}$ each contain $1 + p$ subgroups of type I, of which those generated by U_1U_4 and $U_1^rU_4$ are counted twice. G has therefore $1 + 3p$ subgroups of type I.

Of the groups just described the last contains four independent U's permutable with three subgroups of $\{s_1, s_2\}$, viz., U_1 permutable with s_2, U_2 and U_1U_4 permutable with $s'_1 = s_1s_2^{-1}$, and U_3 permutable with $s_1^{-1}s_2^r$. Hence this group is simply isomorphic with the preceding one. The preceding one itself contains four independent U's permutable with two subgroups of $\{s_1, s_2\}$, viz., U_1 and U_2 permutable with s_2, and U_3 and U_1U_4 permutable with $s'_1 = s_1s_2^{-1}$. This is therefore simply isomorphic with the one which precedes it.

We now suppose that U contains three independent operators of type I. They cannot all be permutable with the same subgroup of $\{s_1, s_2\}$. Suppose first they are permutable with two subgroups. Then generators of G satisfy the relations:

$$
\begin{array}{llll}
U_1^{-1}s_1U_1 = s_1s_3, & U_2^{-1}s_1U_2 = s_1s_4, & U_3^{-1}s_1U_3 = s_1s_5, & \\
U_1^{-1}s_2U_1 = s_2s_j, & & & U_4^{-1}s_2U_4 = s_2s_k.
\end{array}
$$

The group H_1 must be of order p^3 if $\{U_1, U_2, U_3\}$ contains no operators of type I except those in $\{U_2, U_3\}$, and consequently the commutators in the first row may be taken to be s_3, s_4, and s_5. Now s_j is in $\{s_3, s_4, s_5\}$ and consequently $\{U_1, U_2, U_3\}$ does contain operators of type I not in $\{U_2, U_3\}$. Hence the supposition that there are not more than three independent U's of type I and that they are permutable with but two subgroups of $\{s_1, s_2\}$ leads to a contradiction.

We suppose then that the three U's are permutable with three subgroups of $\{s_1, s_2\}$. Generators of G satisfy the relations:

$$
\begin{array}{llll}
U_1^{-1}s_1U_1 = s_1s_j, & U_2^{-1}s_1U_2 = s_1s_3, & U_3^{-1}s_1U_3 = s_1s_4, & \\
U_1^{-1}s_2U_1 = s_2s_k, & & U_3^{-1}s_2U_3 = s_2s_4, & U_4^{-1}s_2U_4 = s_2s_5,
\end{array}
$$

For if s_3 and s_5 were the same, generators of $\{U_2, U_3, U_4\}$ could be selected to give the case above. Now s_j is in $\{s_3, s_4, s_5\}$ and since U_4 may be replaced by any power of itself s_j may be taken to be s_5. Then s_k may be supposed

to be in the group $\{s_3, s_4\}$. Since U_2 may be replaced by any power of itself and U_3 likewise, we may assume that s_k is s_3, s_4, or s_3s_4. If $s_k = s_3$ then the operator $U_1U_2U_3$ is of type I, and U contains four independent operators of type I. If $s_k = s_4$, then $U_1U_3^{-1}$ is of type I. The only possibility is that $s_k = s_3s_4$. The group G exists, it contains but three subgroups of type I, and is therefore distinct from any of those which precede.

When U contains just two independent operators of type I they are permutable with different subgroups of $\{s_1, s_2\}$. Two possibilities arise: (a) U contains $1 + p$ subgroups of type I, or (b) U contains two subgroups of type I.

Let the two independent U's be U_3 and U_4. In case (a) the commutator subgroup of $\{H, U_3, U_4\}$ is of order p; let it be generated by s_5. The commutators arising from transformation of s_1 and s_2 by U_1 may be assumed to be independent of s_5, otherwise $\{U_1, U_3, U_4\}$ would contain operators of type I not in $\{U_3, U_4\}$. These commutators may then be taken to be s_3 and s_4 respectively. Then U_2 may be chosen so that the commutator subgroup of $\{H, U_2\}$ is in $\{s_3, s_4\}$. If $\{U_1, U_2\}$ contains no operator of type I, as must be the case, generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4^r, \quad U_3^{-1}s_1U_3 = s_1s_5,$$
$$U_1^{-1}s_2U_1 = s_2s_4, \quad U_2^{-1}s_2U_2 = s_2s_3, \qquad U_4^{-1}s_2U_4 = s_2s_5.$$

In case (b) we may assume that generators of G satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_5, \quad U_3^{-1}s_1U_3 = s_1s_4,$$
$$U_1^{-1}s_2U_1 = s_2s_j, \quad U_2^{-1}s_2U_2 = s_2s_3, \qquad U_4^{-1}s_2U_4 = s_2s_5.$$

For the commutator subgroup of $\{H, U_1, U_2\}$ is of order at most p^3, and the commutator subgroups of order p^2 arising from transformation of s_1 and s_2 respectively must have an operator in common which does not belong to $\{s_4, s_5\}$. Each of the groups H_1 and H_2 must be of order p^3. Therefore the commutator of U_2 and s_1 can be taken to be s_5. The operator s_j is in $\{s_3, s_4, s_5\}$ and can be assumed to be $s_3^\alpha s_4^\beta$. If α is not zero, then $\{U_1, U_3\}$ contains an operator of type I which is not a power of U_3. We may assume $\alpha = 0$ and $\beta = 1$, in which case $U_1U_2U_3U_4$ is of type I. Therefore there is no group satisfying these conditions.

When $l = 2$, both H_1 and H_2 are of order p^2. It is therefore possible to find generators of G which satisfy the relations:

$$U_1^{-1}s_1U_1 = s_1s_3, \quad U_2^{-1}s_1U_2 = s_1s_4,$$
$$U_3^{-1}s_2U_3 = s_2s_3, \quad U_4^{-1}s_2U_4 = s_2s_4.$$

Hence there is but one group for $l = 2$.

We proceed to list the groups with enough information to determine the group in each case. The first column contains the value of l, which is all the information necessary for the first group and the last. The second column gives the number of subgroups of type I in U. The third column contains whatever further information may be necessary. This additional information in most cases takes the form of a statement of the existence or the non-existence in G of a "subgroup of order p^{n+a} with commutator subgroup of order p^{β} and no operator of type I"; for such a subgroup we shall use the symbol $G_{a,\beta}$. Since it cannot lead to any confusion we shall use the symbol $G_{1,1}$ to stand for a subgroup of type I.

	l	$G_{1,1}$'s	other facts
1.	8	—	—
2.	7	1	—
3.		0	—
4.	6	$1+p$	—
5.		2	—
6.		1	a $G_{3,5}$
7.		1	no $G_{3,5}$
8.		0	a $G_{2,2}$
9.		0	a $G_{3,4}$ and no $G_{2,2}$
10.		0	no $G_{2,2}$, no $G_{3,4}$
11.	5	$1+p+p^2$	—
12.		$2+p$	—
13.		3	—
14.		$1+p$	a $G_{2,3}$
15.		$1+p$	no $G_{2,3}$
16.		$1+p$	[11]
17.		2	a $G_{2,3}$
18.		2	no $G_{2,3}$
19.		1	a $G_{2,2}$
20.		1	a $G_{3,4}$, no $G_{2,2}$
21.		1	no $G_{2,2}$, no $G_{3,4}$
22.	5	0	a $G_{2,2}$
23.		0	a $G_{3,3}$, no $G_{2,2}$
24.		0	no $G_{2,2}$, or $G_{3,3}$
25.	4	$2+p+p^2$	—
26.		$2+2p$	—
27.		$3+p$	—
28.		4	—
	There are $(p+1)/6$ or $(p+5)/6$ such groups.		
29.	4	$1+p$	a $G_{2,2}$
30.		2	a $G_{2,2}$
	There are $(p+1)/2$ such groups.		
31.	4	1	—
	There are $(p+1)/3$ or $(p+5)/3$ such groups.		
32.	4	0	—
	There are $p+2$ such groups. (Cf. above.)		
33.	3	$1+2p+p^2$	—
34.		$1+3p$	—
35.		3	—
36.		$1+p$	—
37.	2	—	—

There are in all $2p+36$ or $2p+38$ distinct groups according as p is of the form $6k-1$ or $6k+1$.

---

[11] Contains a subgroup of order p^{n+2} with commutator subgroup of order p; the two preceding groups have no such subgroup.

2. *The bilinear forms.* As was explained for the case $l = 4$, the commutator structure of G can be described by means of two matrices M and N. In the general case these are matrices of four rows and l columns. That there are two matrices depends on the fact that we are discussing groups G whose centrals are of order p^{n-2}, and therefore generators of H can be chosen so that all but two are in the central of G. The argument in the paper cited above for $l = 4$ holds for any l. A change in the generators of the commutator subgroup of G, when the U's and s_1 and s_2 are not changed, replaces the columns of M by linear combinations of its columns, and replaces the columns of N by the same linear combinations of its columns. The effect is to replace M and N respectively by MB and NB where B is a non-singular square matrix of l rows. A change in the generators of U has the effect of replacing M and N respectively by AM and AN where A is a non-singular four-rowed square matrix. A change in the generators of $\{s_1, s_2\}$ has the effect of replacing the matrix $\lambda_1 M + \lambda_2 N$ by the matrix $(a\lambda'_1 + b\lambda'_2)M + (c\lambda' + d\lambda'_2)N$ where $s'_1 = s_1^a s_2^b$, $s'_2 = s_1^c s_2^d$, and $(ad - bc) \not\equiv 0$.

The matrices M and N may be interpreted as the matrices of two bilinear forms in four variables x_1, x_2, x_3, x_4 and l variables $y_1, y_2, \cdots, y_l$, in which case $\lambda_1 M + \lambda_2 N$ becomes the matrix of a member of the pencil of bilinear forms determined by the two whose matrices are M and N. The changes on generators of the groups then correspond to linear homogeneous non-singular transformations on the x's, the y's, and the λ's. The problem of classification of the groups is then identically the problem of classification of pencils of bilinear forms for $l = 2, \cdots, 8$ under these transformations, for every change of generators of G gives a transformation of the pencil and every transformation of the pencil, with coefficients in the modular field, gives a transformation on generators of the group. In the case of $l = 4$ we were able to classify the groups most easily by means of the theory of invariant factors of $\lambda_1 M + \lambda_2 N$. when M and N are not square a corresponding theory has not been developed. There are some obvious difficulties in the way of extending the theory for square matrices.

Having classified the group for these various values of l we are able to write down immediately a set of normal forms for pencils of bilinear forms in four x's and l y's. We give these normal forms here numbered in the same way as the groups at the end of § 1.

1. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_6 + x_3y_7 + x_4y_8)$.
2. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_6 + x_3y_7)$.
3. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_6 + x_3y_7 + x_4y_1)$.
4. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_6)$.

5. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_2y_5 + x_4y_6)$.
6. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_2y_5 + x_3y_2 + x_4y_6)$.
7. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_2y_5 + x_3y_6 + x_4y_3)$.
8. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_2 + rx_2y_1 + x_3y_5 + x_4y_6)$, r not a square.
9. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_1 + x_3y_2 + x_4y_6)$.
10. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_6 + x_3y_1 + x_4y_2)$.
11. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2x_1y_5$.
12. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_4y_5)$.
13. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_3y_3 + x_4y_5)$.
14. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_1)$.
15. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_3)$.
16. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_2y_5 + x_4y_3)$.
17. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_2y_1 + x_4y_5)$.
18. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_4 + x_2y_3 + x_4y_5)$.
19. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_2 + rx_2y_1 + x_3y_5)$.
20. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_1 + x_3y_2)$.
21. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_1 + x_3y_4)$.
22. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_2 + rx_2y_1 + x_3y_5 + x_4y_3)$.
23. $\lambda_1(x_1y_1 + x_2y_2 + x_3[\alpha y_1 + \beta y_3] + x_4y_4) + \lambda_2(x_1y_3 + x_2y_1 + x_3y_2 + x_4y_5)$.
24. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_5 + x_2y_1 + x_3y_2 + x_4y_3)$.
25. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2x_4y_4$.
26. $\lambda_1(x_1y_1 + x_2y_2) + \lambda_2(x_3y_3 + x_4y_4)$.
27. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_3y_3 + x_4y_4)$.
28. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_2y_2 + \rho x_3y_3 + x_4y_4)$.
29. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_2 + rx_2y_1)$.
30. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_2 + rx_2y_1 + x_4y_4)$.
31. $\lambda_1(x_1y_1 + x_2y_2 + x_3[\alpha y_1 + \beta y_3] + x_4y_4) + \lambda_2(x_1y_3 + x_2y_1 + x_3y_2)$.
32. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + \lambda_2(x_1y_2 + rx_2y_1 + x_3y_4 + rx_4y_3)$.

$$\lambda_1(x_1y_1 + x_2y_2 + x_3y_3 + x_4[\alpha y_4 + \beta y_1 + \gamma y_2 + \delta y_3]) \\ + \lambda_2(x_1y_4 + x_2y_1 + x_3y_2 + x_4y_3),$$

where $\lambda^4 + \delta\lambda^3 - \gamma\lambda^2 + \beta\lambda - \alpha$ has no linear factor.

33. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2x_4y_3$.
34. $\lambda_1(x_1y_1 + x_2y_2) + \lambda_2(x_3y_2 + x_4y_3)$.
35. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1[y_1 + y_3] + x_3y_3 + x_4y_1)$.
36. $\lambda_1(x_1y_1 + x_2y_2 + x_3y_3) + \lambda_2(x_1y_2 + rx_2y_1 + x_4y_3)$.
37. $\lambda_1(x_1y_1 + x_2y_2) + \lambda_2(x_3y_1 + x_4y_2)$.

It is interesting to interpret in terms of the bilinear forms the considera-

tions that were used in classifying the groups. The separation of the groups into classes according to orders of their commutator subgroups corresponds to the separation of the bilinear forms into classes according to the number of the variables y. One of the first things which attracts attention with regard to the groups is that if U is of order p^4 the order of the commutator subgroup cannot be greater than p^8. This says that any pencil of bilinear forms in four x's and l y's is expressible as a pencil of forms in four x's and l' y's where $l' \leq 8$. This is of course obvious from a consideration of the matrices of the forms in terms of which the pencil is expressed.

A pair of numbers λ_1, λ_2 determines a bilinear form of the given pencil; the pair also determines an operator $s_1^{\lambda_1}s_2^{\lambda_2}$ of the group $\{s_1, s_2\}$. When these particular values λ_1, λ_2 are substituted in the expression $\lambda_1 M + \lambda_2 N$ the resulting matrix has four rows each determining an operator of the commutator subgroup arising from transformation of $s_1^{\lambda_1}s_2^{\lambda_2}$ by U. If U contains an operator permutable with $s_1^{\lambda_1}s_2^{\lambda_2}$ the rank of this matrix is less than four. The corresponding bilinear form determined by λ_1, λ_2 will therefore have a rank less than four. Consequently the separation of the groups with commutator subgroups of a given order into classes according to the number of operators of type I in U corresponds to the separation of the pencils of bilinear forms with four x's and l y's into classes according to the number of forms in the pencils which have ranks less than four. We may call such a form, of rank less than four, *singular*. Then the classification of pencils has been made according to the number of singular forms in a given pencil. Theorem (1. 1) states that *any two pencils of bilinear forms in m x's and $2m$ y's are conjugate and that such a pencil contains no singular form.* Theorem (1. 2) states that *there are two distinct pencils of bilinear forms in m x's and $2m - 1$ y's; one contains no singular form, and the other contains one.*[12]

When $l = 6$ the classification according to the number of singular forms in the pencil is not enough. There are two pencils which have each one singular form. In both 6 and 7 above the singular form appears for $\lambda_1, \lambda_2 = 1, 0$. We may distinguish between them in the following way: Consider the numbers (x_1, x_2, x_3, x_4) to be the coördinates of a point in a finite three-space. In terms of the coördinates of the plane $x_4 = 0$ no. 6 determines a pencil of forms in three x's and five y's and no. 7 determines a pencil of forms in three x's and six y's. The first pencil contains no singular form. It is possible to select planes in (x_1, x_2, x_3, x_4) in terms of whose coördinates

[12] Where we understand the forms $F(x_1, \cdots, y_l)$ and $k \cdot F(x_1, \cdots, y_l)$ to be the same; otherwise the number is $p - 1$.

no. 7 will give pencils of forms in three x's and five y's, but every such pencil will contain singular forms.

There are three distinct pencils of forms in four x's and six y's none of which contains a singular form. None of the pencils of forms in the variables of a subspace of (x_1, x_2, x_3, x_4) can be singular. No. 8 determines a pencil of forms in two x's and two y's on the line $x_3 = x_4 = 0$. There is no line on which nos. 9 or 10 determines such a pencil. No. 9 determines a pencil of forms on three x's and four y's on the plane $x_4 = 0$, and there is no plane on which no. 10 determines such a pencil.

The interpretation in terms of bilinear forms is particularly enlightening in the case of no. 28. There are $(p+1)/6$ or $(p+5)/6$ such pencils depending on the value of ρ. Each pencil contains four singular forms, each singular form is given by a pair λ_1, λ_2. Each form of the pencil determines a point on the line (λ_1, λ_2). The cross-ratio of these four points remains invariant under projective transformation of the ordered set of four points. A reordering of the four points gives one of six values of the cross-ratio. Two forms determined by ρ and ρ' cannot be conjugate unless ρ' is one of the values ρ, $1-\rho$, $1/\rho$, $1/(1-\rho)$, $(\rho-1)/\rho$, or $\rho/(\rho-1)$.

The differences among the groups that come under 30, 31, and 32 can all be interpreted in terms of pencils induced in subspaces of (x_1, x_2, x_3, x_4). It is suggested that perhaps a more thoroughgoing geometric interpretation of the whole situation would be worth while.

3. *The proof of distinctness.* We come now to the question of the possible isomorphism of two groups belonging to different classes. Our proofs of uniqueness of the various normal forms of generating relations have always been proofs of uniqueness under automorphisms of G in which H corresponds to itself. In none of these groups has H been the only abelian group of order p^n in G, for the group $\{U_1, U_2, s_3, s_4, \cdots, s_n\}$ is such a group. Moreover, the group U is not in general a characteristic subgroup of G, for any operator U_i may be replaced by sU_i where s is any operator of H. If s is in the central of G this has no effect on generating relations, but if s is not in the central the U's will in general cease to be permutable. We are confining our attention to groups G in which the U's are permutable. This is justified by the fact that any classification of metabelian groups must take these groups into account; it must depend on the possibilities of "commutator structure" arising from transformation of H by U. The simplicity of the statement in terms of pencils of bilinear forms gives added assurance that the limitations imposed on the investigations are natural to the problem and not arbitrary.

The question may be considered from the point of view of the uniqueness of the defining relations. The particular defining relations that we have chosen in each case are not the only ones that could be chosen. The situation is rather that if certain properties of the generators are required, then the defining relations can be reduced to the normal form we have given for the particular group. If two groups G and G' belonging to different classes were simply isomorphic, then it would be possible to select generators of G' which satisfied defining relations of G. These generators would in every case satisfy the following conditions:

(1) n of the generators would generate an abelian subgroup H', and and $n-2$ of them would generate the central of G';
(2) the remaining four would generate an abelian group of type $1, 1, 1, 1$;
(3) no operator of the second group, except identity, would be permutable with every operator of the first;
(4) the operators of the second group would correspond to operators of types I and II in the I-group of the first.

We have already considered the possibility of such isomorphisms between G and G' in which H corresponds to itself and have seen that none exists when the normal forms of the generating relations are different. The group H' cannot then be $\{s_1, s_2, \cdots, s_n\}$; it must, however, include the central $\{s_3, s_4, \cdots, s_n\}$. We assume then that G and G' have different normal forms for their generating relations and that there exists a simple isomorphism between them. Then there exists in G a group H' which corresponds to H in G'. This group H' must be $\{s'_1, s'_2, s_3, s_4, \cdots, s_n\}$, where

$$s'_i = s_1^{\alpha_i} s_2^{\beta_i} U_1^{k_i} U_2^{l_i} U_3^{m_i} U_4^{n_i}, \qquad (i = 1, 2).$$

The numbers $k_1, \cdots, n_2$ cannot all be zero for then H' would be H. We may suppose that all but k_1 and l_2 or all but k_1 and k_2 are zero, for this requires only a proper choice of generators of U. We may suppose further that the two which are not zero are both equal to 1, so that $s'_1 = s_1^{\alpha_1} s_2^{\beta_1} U_1$ and $s'_2 = s_1^{\alpha_2} s_2^{\beta_2} U_i$, $(i = 1 \text{ or } 2)$. If $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$ we may select generators of $\{s_1, s_2\}$ so that $s'_1 = s_1 U_1$ and $s'_2 = s_2 U_i$. If then $i = 1$, we may replace s'_2 by $s_2'' = s_2$. If $i = 2$, the group $U' = \{U'_1, U'_2, U'_3, U'_4\}$ where

$$U'_i = s_1^{\mu_i} s_2^{\nu_i} U_1^{k_i} U_2^{l_i} U_3^{m_i} U_4^{n_i}$$

must contain U_1 or s_1 and U_2 or s_2. U' cannot contain both s_1 and s_2 for then without changing generating relations we could replace H' by

$H'' = \{U_1, U_2, s_3, s_4, \cdots, s_n\}$ which is not maximal abelian invariant. We may therefore suppose that $s'_1 = s_1U_1$ and $s'_2 = s_2$. If on the other hand $\alpha_1\beta_2 - \alpha_2\beta_1 = 0$, we may suppose that $s'_1 = s_1U_1$ and $s'_2 = U_i$. If $U_i = U_1$, s_1 is in H' and we may take s'_1 to be s_1. If U_i is not U_1, then either U_1 or s_1 is in U'. We have seen that there are not two independent U's in H' and therefore whatever the value of $\alpha_1\beta_2 - \alpha_2\beta_1$ we may suppose that

$$s'_1 = s_1U_1 \quad \text{and} \quad s'_2 = s_2.$$

From this it follows that U_1 must be permutable with s_2 and hence must be of type I. In the expression for one of the U'_i's as given above in terms of $s_1, s_2, U_1, \cdots, U_4$ the exponent μ_i must be different from zero. Making use of it, s'_1 may be replaced by $s_1'' = U_1$ without affecting the generating relations. The group U' must then contain s_1 or else by the same sort of transformation H' can be changed to H. But if s_1 is in U' it must be permutable with U_2, U_3, and U_4. This identifies the group as no. 25. Since this is the only one of the groups with $2 + p + p^2$ subgroups of type I, the original sets of relations of generators of G and G' were transformable into each other.

4. *The general case.* The theorems (1.1), (1.2), and (1.4) go beyond the case where the order of U is p^4. It is clear that the methods used will suffice to classify the groups G of order p^{n+m} where U of order p^m and abelian of type $1, 1, \cdots$ contains only operators of types I and II. It is clear also that this problem is the same as the problem of classification of families of forms. The groups of order p^{n+m} may be separated first into classes according to the orders of their commutator subgroups. Then each of these classes may be subdivided according to the number of independent operators of type I in U. When these U's of type I are segregated, the remaining operators of a set of independent generators of U determine a group U' whose operators are all of type II. U' determines two groups H'_1 and H'_2 which are commutator subgroups arising from transformation of s_1 and s_2 respectively by U'. Unless the commutator subgroup of $\{H, U'\}$ is the product of the two distinct groups H'_1 and H'_2, H'_1 and H'_2 have a cross-cut different from identity. This cross-cut determines operators $U_1, U_2, \cdots$ such that their commutators with s_1 generate the cross-cut and it determines operators $V_1, V_2, \cdots$ such that their commutators with s_2 generate the cross-cut. The order of the group $U'' = \{U_1, U_2, \cdots, V_1, V_2, \cdots\}$ enables us to determine a normal form for the relations defining $\{H, U''\}$ and then a normal form for relations defining G. The kinds of differences that may present themselves are apparent. For example, the operators $U_1, U_2, \cdots, V_1, V_2, \cdots$ may be independent or they

may be dependent in various ways. It may be possible to select α U's and α V's such that the commutator subgroup determined by them and H is of order p^{α}, in which case G contains a subgroup $G_{\alpha,\alpha}$. If that is the case it is necessary to determine the type of $G_{\alpha,\alpha}$ which appears. This goes back to the question of the invariant factors of the matrix $\lambda_1 M + \lambda_2 N$ where M and N are α-rowed square matrices. Though the method is clear and obviously sufficient it would be desirable to continue the study on account of the interesting facts that are bound to appear in the classification of polynomials of degree even as small as five.

It is likewise obvious that the methods used here are sufficient for the classification of groups where U contains operators of type more "advanced" than I and II. If U contains an operator of type III, one which determines the partition $n = 2 + 2 + 2 + 1 + \cdots + 1$, and no operator except those of types I, II, and III, then the central of G would be of order p^{n-3}. Only three of the generators of H, s_1, s_2, and s_3, need be outside the central of G. They would determine three matrices M_1, M_2, M_3 and three bilinear forms. The classes of groups would correspond in a $1 - 1$ manner with the classes of three-parameter (homogeneous) families of bilinear forms $\lambda_1 M_1 + \lambda_2 M_2 + \lambda_3 M_3$. This classification would probably depend on the types of pencil as well as the types of form contained in the family. The method of procedure is clear in its general aspects; the details of the possible difficulties are not so clear. On that account the classification should be carried on in detail somewhat further in this direction.

The methods and results point the way also to the treatment of groups where U contains operators of type $\bar{K}$ and none of type greater than $\bar{K}$. We shall content ourselves with the statement of the following theorem which has been clear for some time, although the theorem does not take full account of the method of attack we have used.

The problem of the classification of metabelian groups $\{H, U\}$ of order p^{n+m} which contain H as a maximal invariant abelian group and in which U is in the group of isomorphisms of H is identical with the problem of classification of k-parameter (homogeneous) families of bilinear forms in m variables x and an undetermined number of variables y under "rational" projective transformation on the x's, the y's, and on the parameters. The number k takes on all values not greater than $n/2$.

UNIVERSITY OF ILLINOIS.

A METRICALLY TRANSITIVE GROUP DEFINED BY THE MODULAR GROUP.

By Gustav A. Hedlund.[1]

1. Introduction. It is known [2] that the modular group, Γ, is metrically transitive with respect to the real axis. This means that if E is a measurable point set of the real axis which is invariant under the transformations of the modular group, Γ, either E or its complement with respect to the real axis is a zero set.

Let Γ_2 be the group of real transformations of the (ξ, η) plane, π, into itself, given by

$$T: \qquad \xi = \frac{a\xi' + b}{c\xi' + d}, \qquad \eta = \frac{a\eta' + b}{c\eta' + d},$$

where a, b, c and d are integers such that $ad - bc = 1$. The object of this paper is to prove that the group Γ_2 is metrically transitive with respect to the plane π.

This is the essential result needed in proving the metrical transitivity of the dynamical system obtained by considering the non-euclidean billiard problem [3] defined by the modular group. The similar result where the modular group is replaced by a certain Fuchsian group with closed fundamental region has been published by the author.[4] The technic in the two cases is similar, but in the present case the method is not buried under the details involved in the other case.

The result obtained here implies the metrical transitivity of the group Γ with respect to the real axis, but the proof is, of course, very indirect. Conversely, as an example shows, metrical transitivity of a group G with respect to the real axis does not necessarily imply metrical transitivity of the group G_2, which is obtained by applying the transformations of G simultaneously to two variables, with respect to the plane of the two variables.

---

[1] This paper was completed during the tenure of a National Research Fellowship.

[2] M. H. Martin, "Metrically Transitive Point Transformations," *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 606-612.

[3] E. Artin, "Ein mechanisches System mit quasiergodischen Bahnen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 3 (1924), pp. 170-175.

[4] G. A. Hedlund, "On the metrical transitivity of the geodesics on closed surfaces of constant negative curvature," *Annals of Mathematics*, vol. 35 (1934), pp. 787-808.

2. Quasi-transitive points. Under the transformations of the group Γ_2 a point P of π is transformed into the set of points congruent to P. If this set is everywhere dense in π the point P will be called a *quasi-transitive* point. From the work of Artin [5] it is known that not only are there quasi-transitive points, but a point is quasi-transitive if either its abscissa or ordinate belongs to a certain linear set, the complement of which with respect to the real axis is a linear zero set. It immediately follows that the set of non-quasi-transitive points of π constitutes a zero set.

With the aid of this result the problem of proving that any invariant measurable set in π is either a zero set or the complement of a zero set is considerably simplified. For since the non-quasi-transitive points form an invariant zero set, these points can be omitted from a measurable invariant set without affecting either the measure or the invariance. The use of this fact is illustrated in the following lemma.

LEMMA 2. 1. *If E is an invariant measurable set of π and D is an open set of π, E is a zero set if $E \cdot D$ is a zero set.*

For assuming all points of E quasi-transitive, the set E can be obtained from the set $E \cdot D$ by applying the transformations of the group Γ_2. But if $E \cdot D$ is a zero set, the set obtained by applying the denumerable set of transformations of Γ_2 is a zero set.

In particular the set D will be chosen as the set $1 < \xi < 2, -1 < \eta < 0$. If E is the given invariant measurable set and it is shown that either $E \cdot D$ or $D - E \cdot D$ is a zero set, the desired theorem will have been proved.

3. A net in D. Let $P(\xi, \eta)$ be a point of D and let the developments of ξ and $-\eta$ in continued fractions with positive integral partial quotients be given by

$$\xi = [1, a_1, a_2, \cdots], \qquad -\eta = [0, a_{-1}, a_{-2}, \cdots],$$

where $[b_0, b_1, b_2, \cdots]$ is given by [6]

$$[b_0, b_1, \cdots] = b_0 + \cfrac{1}{b_1 + \cfrac{1}{b_2 + \text{etc.},}}$$

and the continued fraction may or may not be terminating. If it is terminating, the representation will not be unique, but this does not affect the following arguments.

[5] E. Artin, *loc. cit.*, p. 174.

[6] Perron, *Die Lehre von den Kettenbrüchen*, Teubner, 1913, p. 27.

Given the two sets of positive integers, $a_1, a_2, \cdots, a_\mu$, and $a_{-1}, a_{-2}, \cdots, a_{-\nu}$, let R be those points (ξ, η) of D such that $\xi = [1, a_1, a_2, \cdots, a_\mu, \cdots]$ and $-\eta = [0, a_{-1}, a_{-2}, \cdots, a_{-\nu}, \cdots]$, where again these may be terminating continued fractions but they must contain enough partial quotients to assure the presence of the given sequences. The set R includes all the points of a rectangle in D. The interior, Δ, of R, will be denoted by

$$\{1, a_1, \cdots, a_\mu;\ 0, a_{-1}, \cdots, a_{-\nu}\}.$$

Thus, in particular, D is given by $\{1; 0\}$.

Let $L(a_1, a_2, \cdots, a_\mu)$ be the length of a side of Δ parallel to the x-axis and $L(a_{-1}, a_{-2}, \cdots, a_{-\nu})$ the length of a vertical side. The following formula is readily obtained:

$$(3.1) \qquad L(a_1, \cdots, a_\mu) = \prod_{i=1}^{\mu} (P_i P'_i)^{-1},$$

where $P_i = [a_i, \cdots, a_\mu]$ and $P'_i = [a_i, \cdots, a_{\mu-1}, a_\mu + 1]$, $(i = 1, 2, \cdots, \mu)$. A similar formula holds for $L(a_{-1}, \cdots, a_{-\nu})$.

4. A lemma on continued fractions.

LEMMA 4.1.

$$F(x_0, x_1, \cdots, x_n) = \frac{[x_0, x_2, \cdots, x_n]}{[x_0, x_1, \cdots, x_n + \alpha]}, \quad 0 \leqq \alpha,\ 1 \leqq x_i < \infty.$$

If n is even (odd), F is a non-decreasing (non-increasing) function of each of the variables x_i, $(i = 0, 1, \cdots, n)$.

Let $P_i = [x_i, \cdots, x_n]$ and $Q_i = [x_i, \cdots, x_{n-1}, x_n + \alpha]$. Then obviously, $\partial P_0/\partial x_0 = \partial Q_0/\partial x_0 = 1$, and the following formulas can be obtained by evaluating the limit of the difference quotient:

$$(4.2) \quad \frac{\partial P_0}{\partial x_i} = (-1)^i \prod_{k=1}^{i} P_k^{-2}; \quad \frac{\partial Q_0}{\partial x_0} = (-1)^i \prod_{k=1}^{i} Q_k^{-2}; \quad (i = 1, 2, \cdots, n).$$

From these formulas follow:

$$(4.3) \quad \frac{\partial F}{\partial x_0} = \frac{Q_0 - P_0}{Q_0^2}; \quad \frac{\partial F}{\partial x_i} = (-1)^i Q_0^{-1} \left\{ 1 - \frac{P_0}{Q_0} \prod_{k=1}^{i} \frac{P_k^2}{Q_k^2} \right\} \prod_{k=1}^{i} P_k^2, \quad (i = 1, 2, \cdots, n).$$

Case I, n even. Then $P_i \leqq Q_i$, i even, and $P_i \geqq Q_i$, i odd. In this case it follows at once from (4.3) that $\partial F/\partial x_0 \geqq 0$. From (4.3) we have

$$\frac{\partial F}{\partial x_1} = -Q_0^{-1} P_1^{-2} \left\{ 1 - \frac{P_0}{Q_0} \frac{P_1^2}{Q_1^2} \right\}.$$

But

$$\frac{P_0P_1}{Q_0Q_1}=\frac{x_0P_1+1}{x_0Q_1+1}\geqq 1,$$

since $P_1 \geqq Q_1$, and hence $\partial F/\partial x_1 \geqq 0$. To complete the proof by induction, we assume the theorem true for $(i=1,2,\cdots,j<n)$. Then if j is even we wish to show that $P_0Q_0^{-1}\prod_{k=1}^{j+1}P_k{}^2Q_k{}^{-2}\geqq 1$. From the assumption that the result holds for $i<j+1$, it follows that $P_0Q_0^{-1}\prod_{k=1}^{j-1}P_k{}^2Q_k{}^{-2}\geqq 1$. But the inequality $P_{j+1}\geqq Q_{j+1}$ implies that $P_jP_{j+1}/Q_jQ_{j+1}=(x_jP_{j+1}+1)/(x_jQ_{j+1}+1)\geqq 1$, and the desired result follows. The case where j is odd is treated similarly.

Case II, n odd. Then $P_i \geqq Q_i$, i even, $P_i \leqq Q_i$, i odd, and from the reversal of the inequalities, the proof in this case follows readily from that given in Case I.

5. The fundamental lemma.

LEMMA 5.1. *There exist positive constants k_1 and k_2 such that*

$$k_1\frac{L(a'_1,\cdots,a'_m,a_{m+1},\cdots,a_n)}{L(a'_1,\cdots,a'_m)}<\frac{L(a_1,\cdots,a_m,a_{m+1},\cdots,a_n)}{L(a_1,\cdots,a_m)}$$
$$<k_2\frac{L(a'_1,\cdots,a'_m,a_{m+1},\cdots,a_n)}{L(a'_1,\cdots,a'_m)},\qquad n\geqq m>0,$$

independent of the positive integers $a'_1,\cdots,a'_m,a_1,\cdots,a_m,a_{m+1},\cdots,a_n$.

Proof. Let

$$\begin{aligned}
&P_i=[a_i,\cdots,a_n]; && P'_i=[a_i,\cdots,a_n+1], \quad (i=1,\cdots,n);\\
&Q_i=[a'_i,\cdots,a'_m,a_{m+1},\cdots,a_n]; && Q'_i=[a'_i,\cdots,a'_m,a_{m+1},\cdots,a_n+1],\\
& && \qquad (i=1,\cdots,m);\\
&R_i=[a_i,\cdots,a_m]; && R'_i=[a_i,\cdots,a_m+1], \quad (i=1,\cdots,m);\\
&S_i=[a'_i,\cdots,a'_m]; && S'_i=[a'_i,\cdots,a'_m+1], (i=1,\cdots,m).
\end{aligned}$$

Then

$$\begin{aligned}
L(a_1,\cdots,a_m,a_{m+1},\cdots,a_n)&=\prod_{i=1}^{n}(P_iP'_i)^{-1};\\
L(a'_1,\cdots,a'_m,a_{m+1},\cdots,a_n)&=\prod_{i=1}^{m}(Q_iQ'_i)^{-1}\prod_{i=m+1}^{n}(P_iP'_i)^{-1};\\
L(a_1,\cdots,a_m)&=\prod_{i=1}^{m}(R_iR'_i)^{-1};\\
L(a'_1,\cdots,a'_m)&=\prod_{i=1}^{m}(S_iS'_i)^{-1};
\end{aligned}$$

and

$$\frac{L(a_1,\cdots,a_n)L(a'_1,\cdots,a'_m)}{L(a_1,\cdots,a_m)L(a'_1,\cdots,a'_m,a_{m+1},\cdots,a_n)}=\prod_{i=1}^{m}\frac{R_iR'_iQ_iQ'_i}{P_iP'_iS_iS'_i}$$
$$=G(a_1,\cdots,a_m,a'_1,\cdots,a'_m,a_{m+1},\cdots,a_n).$$

Let us consider

$$\frac{R_i}{P_i}=\frac{[a_i,\cdots,a_m]}{[a_i,\cdots,a_m,a_{m+1},\cdots,a_n]}.$$

From Lemma 4. 1, if $m-i$ is even,

$$\frac{R_i(1)}{R_{i-1}(1)}\leqq\frac{R_i}{P_i}\leqq 1,$$

where $R_i(1)$ is the number obtained by replacing each argument in R_i by 1. Similarly, if $m-i$ is odd

$$1\leqq\frac{R_i}{P_i}\leqq\frac{R_i(1)}{R_{i-1}(1)}.$$

Hence

$$\left(\frac{R_m(1)}{R_{m-1}(1)}\cdot\frac{R_{m-2}(1)}{R_{m-3}(1)}\cdots\right)\prod_{i=1}^{m}\frac{R'_iQ_iQ'_i}{P'_iS_iS'_i}$$
$$\leqq G\leqq\left(\frac{R_{m-1}(1)}{R_{m-2}(1)}\,\frac{R_{m-3}(1)}{R_{m-4}(1)}\cdots\right)\prod_{i=1}^{m}\frac{R'_iQ_iQ'_i}{P'_iS_iS'_i}$$

where the products in the parentheses are continued as long as the subscripts remain positive. From this follows

$$\left(\frac{[1]}{[1,1]}\,\frac{[1,1,1]}{[1,1,1,1]}\cdots\right)\prod_{i=1}^{m}\frac{R'_iQ_iQ'_i}{P'_iS_iS'_i}$$
$$\leqq G\leqq\left(\frac{[1,1]}{[1,1,1]}\,\frac{[1,1,1,1]}{[1,1,1,1,1]}\cdots\right)\prod_{i=1}^{m}\frac{R'_iQ_iQ'_i}{P'_iS_iS'_i},$$

provided the infinite products in the parentheses converge. But the successive convergents of $(1+5^{1/2})/2$ are given by $c_1=[1]$, $c_2=[1,1],\cdots$, and it is readily shown that the infinite products $\prod_{i=1}^{\infty}c_{2i-1}/c_{2i}$ and $\prod_{i=1}^{\infty}c_{2i}/c_{2i+1}$ converge to positive constants e_1 and e_2, respectively.

Using the same technic it can be shown that the desired inequality

$$e_1{}^2/e_2{}^2\leqq G\leqq e_2{}^2/e_1{}^2$$

obtains. Setting $k_1=e_1{}^2/e_2{}^2$ and $k_2=e_2{}^2/e_1{}^2=k_1{}^{-1}$, the desired lemma holds.

6. Some inequalities. Let $\Delta=\{1,a_1,\cdots,a_\mu;\ 0,a_{-1},a_{-2},\cdots,a_{-\nu}\}$ be a chosen rectangle of the net in D. Let σ_λ denote the set of sub-rectangles

of Δ given by $\{1, a_1, \cdots, a_\mu, s_1, s_2, \cdots, s_\lambda, 1;\ 0, a_{-1}, a_{-2}, \cdots, a_{-\nu}\}$, where $s_1, s_2, \cdots, s_\lambda$ are arbitrary positive integers, but λ is so chosen that $\lambda + \mu$ is odd. The area of the set σ_λ is denoted by $\mu(\sigma_\lambda)$.

LEMMA 6.1. *There exists a positive constant k_3 such that*

$$\mu(\sigma_\lambda) > k_3\mu(\Delta).$$

This amounts simply to proving that there exists a positive constant k_3 such that

$$L(a_1, a_2, \cdots, a_m, 1) > k_3 L(a_1, a_2, \cdots, a_m),$$

where k_3 is independent of $a_1, a_2, \cdots, a_m$.

A brief computation with the aid of (3.1) yields

$$\frac{L(a_1, a_2, \cdots, a_m, 1)}{L(a_1, a_2, \cdots, a_m)} = \frac{[a_1, a_2, \cdots, a_m]}{[a_1, a_2, \cdots, a_m + 1/2]} \cdot \frac{[a_2, \cdots, a_m]}{[a_2, \cdots, a_m + 1/2]} \cdots \frac{a_m}{a_m + 1/2} \cdot \frac{1}{2}.$$

With the aid of Lemma 4.1 and using the notation of the preceding paragraph,

$$\frac{L(a_1, a_2, \cdots, a_m, 1)}{L(a_1, a_2, \cdots, a_m)} \geqq \frac{1}{c_2} \cdot \frac{1}{c_3} \cdot \frac{c_3}{c_5} \cdot \frac{c_5}{c_7} \cdots = \frac{1}{1 + 5^{1/2}}$$

and the above lemma holds with $k_3 = (1 + 5^{1/2})$.

Using the fact that $\lambda + \mu$ is odd, the rectangle

$$\bar{\Delta} = \{1, a_1, a_2, \cdots, a_\mu, s_1, \cdots, s_\lambda, 1;\ 0, a_{-1}, a_{-2}, \cdots, a_{-\nu}\}$$

is defined by the inequalities:

$$\begin{aligned} \xi_1 &= [1, a_1, \cdots, a_\mu, s_1, \cdots, s_\lambda, 1] < \xi < [1, a_1, \cdots, a_\mu, s_1, \cdots, s_\lambda, 2] = \xi_2, \\ -\eta_1 &= [0, a_{-1}, \cdots, a_{-\nu}] \qquad\qquad < -\eta < [0, a_{-1}, \cdots, a_{-\nu} + 1] \qquad = -\eta_2, \end{aligned}$$

provided ν is odd. The second inequality is reversed if ν is even, but it will be sufficient to discuss the case ν odd.

The transformation $\xi = [1, a_1, \cdots, a_\mu, s_1, \cdots, s_\lambda, \xi']$ is given by

$$\xi = \frac{a\xi' + b}{c\xi' + d}, \qquad ad - bc = (-1)^{\lambda+\mu+1} = 1, \tag{6.1}$$

where a, b, c and d are positive integers and hence (6.1) is a member of the group Γ. If we let

$$-\eta'_1 = [0, s_\lambda, s_{\lambda-1}, \cdots, s_1, a_\mu, \cdots, a_1, 1, a_{-1}, \cdots, a_{-\nu}]$$

and

$$-\eta'_2 = [0, s_\lambda, s_{\lambda-1}, \cdots, s_1, a_\mu, \cdots, a_1, 1, a_{-1}, \cdots, a_{-\nu} + 1],$$

then it is easily shown (this result is stated in Artin, *loc. cit.*, p. 173) that the relations,

$$\eta_1 = \frac{a\eta' + b}{c\eta' + d}, \qquad \eta_2 = \frac{a\eta'_2 + b}{c\eta'_2 + d},$$

hold, where a, b, c and d are given in (6. 1). Thus it is seen that there is a transformation of the group Γ_2 which transforms the rectangle

$$\bar{\Delta} = \{1, a_1, \cdots, a_\mu, s_1, \cdots, s_\lambda, 1;\ 0, a_{-1}, a_{-2}, \cdots, a_{-\nu}\}$$

into the rectangle

$$\bar{\Delta}' = \{1;\ 0, s_\lambda, s_{\lambda-1}, \cdots, s_1, a_\mu, \cdots, a_1, 1, a_{-1}, \cdots, a_{-\nu}\}.$$

Using only transformations of the group Γ_2, each of the non-overlapping rectangles of σ_λ, for fixed λ, can be transformed into one of the non-overlapping set σ'_λ, $\{1; s_\lambda, \cdots, s_1, a_\mu, \cdots, a_1, 1, a_{-1}, \cdots, a_{-\nu}\}$, such that the correspondence is one-to-one.

LEMMA 6. 2. *There exists a positive constant k_4 such that*

$$\mu(\sigma'_\lambda) > k_4\mu(\sigma_\lambda).$$

Considering a single pair, $\bar{\Delta}$ and $\bar{\Delta}'$, of corresponding rectangles of the sets σ_λ and σ'_λ,

$$\mu(\Delta) = |\xi_2 - \xi_1| \cdot |\eta_2 - \eta_1|,$$

and

$$\mu(\bar{\Delta}') = |\eta'_2 - \eta'_1|.$$

A brief computation yields the equalities

$$\frac{\mu(\bar{\Delta}')}{\mu(\bar{\Delta})} = \frac{|\eta'_2 - \eta'_1|}{|\xi_2 - \xi_1| \cdot |\eta_2 - \eta_1|}$$
$$= |(c\xi'_1 + d)(c\xi'_2 + d)(c\eta'_1 + d)(c\eta'_2 + d)| = \frac{|(\xi'_1 - \eta'_1)(\xi'_2 - \eta'_2)|}{|(\xi_1 - \eta_1)(\xi_2 - \eta_2)|}.$$

From the inequalities $1 \leqq \xi_1, \xi'_1, \xi_2, \xi'_2 \leqq 2$, $-1 \leqq \eta_1, \eta'_1, \eta_2, \eta'_2 \leqq 0$, it follows from the last equation that

$$\frac{\mu(\bar{\Delta}')}{\mu(\bar{\Delta})} \geqq \frac{1}{9}.$$

But this inequality holds for each of the corresponding pairs in σ_λ and σ'_λ and the desired lemma holds with $k_4 = 1/9$.

LEMMA 6.3. *If E is a measurable point set of π which is invariant under the group Γ_2, then*

$$\mu(E\cdot\sigma_\lambda) > k_4\mu(E\cdot\sigma'_\lambda).$$

Denoting, as above, by $\bar{\Delta}$ and $\bar{\Delta}'$ corresponding rectangles of the sets σ_λ and σ'_λ respectively, since E is an invariant set, the set $E\cdot\bar{\Delta}$ is transformed into the set $E\cdot\bar{\Delta}'$ by the transformation of the group Γ_2 taking $\bar{\Delta}$ into $\bar{\Delta}'$. Using the equation

$$\frac{\mu(\bar{\Delta})}{\mu(\bar{\Delta}')} = \frac{|(\xi_1-\eta_1)(\xi_2-\eta_2)|}{|(\xi'_1-\eta'_1)(\xi'_2-\eta'_2)|}$$

there is obtained as above

$$\mu(\bar{\Delta}) \geqq k_4\mu(\bar{\Delta}').$$

But if $\bar{\bar{\Delta}}$ is any sub-rectangle, with sides parallel to the axes, of $\bar{\Delta}$, and $\bar{\bar{\Delta}}'$ is the corresponding rectangle under T, precisely the same inequality can be obtained, viz.,

$$\mu(\bar{\bar{\Delta}}) \geqq k_4\mu(\bar{\bar{\Delta}}').$$

The rectangle $\bar{\bar{\Delta}}$ being an arbitrary sub-rectangle (with sides parallel to the axes) of $\bar{\Delta}$, it follows that the Jacobian of T, evaluated at any point of $\bar{\Delta}$, lies between 9 and 1/9. This implies the inequality $\mu(E\cdot\bar{\Delta}) \geqq k_4\mu(E\cdot\bar{\Delta}')$. By summing over the set σ_λ, the stated lemma is obtained.

LEMMA 6.4. *E being a measurable invariant set of π, there exists a positive constant, k_5, which does not depend on how Δ is chosen in D, and is such that if λ is chosen sufficiently large,*

$$\mu(E\cdot\sigma'_\lambda) \geqq k_5\mu(E\cdot D)\mu(\sigma'_\lambda).$$

Case I. $E\cdot D = \{1, e_1, e_2, \cdots, e_m;\ 0, e_{-1}, e_{-2}, \cdots, e_{-n}\}$.

Let λ be so chosen that $\lambda > n$. Then

$$\mu(E\cdot\sigma'_\lambda) = L(e_1,\cdots,e_m)\sum_{s_{\lambda-n}\cdots s_1} L(e_{-1},\cdots,e_{-n},s_{\lambda-n},\cdots,s_1,a_\mu,\cdots,a_1,1,a_{-1},\cdots,a_{-\nu}),$$

where $\sum\limits_{s_{\lambda-n}\cdots s_1}$ indicates the sum of the lengths of the intervals for which $s_{\lambda-n},\cdots,s_1$ are arbitrary positive integers, but the other elements are fixed. Similarly,

$$\mu(\sigma'_\lambda) = \sum_{s_\lambda\cdots s_1} L(s_\lambda, s_{\lambda-1},\cdots,s_1,a_\mu,\cdots,a_1,1,a_{-1},\cdots,a_{-\nu}),$$

where the sum is extended over all positive integral values of $s_\lambda,\cdots,s_1$.

From Lemma 5.1,

$$k_1 \frac{L(1,\cdots,1,a_\mu,\cdots,a_{-\nu})}{L(1,\cdots,1)} < \frac{L(s_\lambda,s_{\lambda-1},\cdots,s_1,a_\mu,\cdots,a_{-\nu})}{L(s_\lambda,\cdots,s_1)} < k_2 \frac{L(1,\cdots,1,a_\mu,\cdots,a_{-\nu})}{L(1,\cdots,1)}$$

where in each term where $1,\cdots,1$ occurs, there are λ such elements. It follows from this that

$$\sum_{s_\lambda\cdots s_1} L(s_\lambda,\cdots,s_1,a_\mu,\cdots,a_{-\nu}) < k_2 \frac{L(1,\cdots,1,a_\mu,\cdots,a_{-\nu})}{L(1,\cdots,1)} \sum_{s_\lambda\cdots s_1} L(s_\lambda,\cdots,s_1).$$

Evidently

$$\sum_{s_\lambda\cdots s_1} L(s_\lambda,\cdots,s_1) = 1,$$

and

$$\sum_{s_{\lambda-n}\cdots s_1} L(e_{-1},e_{-2},\cdots,e_{-n},s_{\lambda-n},\cdots,s_1) = L(e_{-1},e_{-2},\cdots,e_{-n}).$$

With the aid of these

$$\frac{\mu(E\cdot\sigma'_\lambda)}{\mu(\sigma'_\lambda)} \geqq k_1k_2^{-1}L(e_{-1},e_{-2},\cdots,e_{-n})L(e_1,\cdots,e_m) = k_1k_2^{-1}\mu(E\cdot D).$$

Choosing $k_5 = k_1/k_2$, the lemma holds in this case.

Case II. $E\cdot D$ is a finite set of non-overlapping rectangles of the net in D.

Let $E\cdot D = \sum_{i=1}^{N} R_i$. If λ is chosen sufficiently large the proof given in Case I holds simultaneously for all of the set R_i, $(i=1,\cdots,N)$, and hence

$$\mu(E\cdot\sigma'_\lambda) = \sum_{i=1}^{N}\mu(R_i\cdot\sigma'_\lambda) \geqq \sum_{i=1}^{N} k_1k_2^{-1}\mu(R_i)\mu(\sigma'_\lambda) = k_1k_2^{-1}\mu(\sigma'_\lambda)\mu(E\cdot D).$$

The lemma holds again with $k_5 = k_1k_2^{-1}$.

Case III. $E\cdot D$ is an infinite set of non-overlapping rectangles of the net in D; $E\cdot D = \sum_{i=1}^{\infty} R_i$.

Given ϵ, there exists an N such that $\mu(\sum_{i=1}^{N} R_i) > (1-\epsilon)\mu(E\cdot D)$. For any λ, $\mu(E\cdot\sigma'_\lambda) = \sum_{i=1}^{\infty}\mu(R_i\cdot\sigma'_\lambda) \geqq \sum_{i=1}^{N}\mu(R_i\cdot\sigma'_\lambda)$. If λ is sufficiently large, Case II can be applied and

$$\mu(E\cdot\sigma'_\lambda) \geqq \sum_{i=1}^{N}\mu(R_i\cdot\sigma'_\lambda) \geqq k_1k_2^{-1}\mu(\sigma'_\lambda)\sum_{i=1}^{N}\mu(R_i) > k_1k_2^{-1}(1-\epsilon)\mu(E\cdot D)\mu(\sigma'_\lambda).$$

If $\epsilon < 1/3$, the lemma holds with $k_5 = 2k_1/3k_2$.

Case IV. $E \cdot D$ is an open set. This case is already included under III, for an open set is the sum of an infinite set of non-overlapping rectangles of the net in D, together with the boundaries of these rectangles. Since the boundaries form a zero set, they do not affect the argument.

Case V. $E \cdot D$ is a measurable set of positive measure.

Given $a_1, \cdots, a_\mu, a_{-1}, \cdots, a_{-\nu}$, from Lemma 5. 1,

$$\begin{aligned}\mu(\sigma'_\lambda) &= \sum_{s_\lambda \cdots s_1} L(s_\lambda, \cdots, s_1, a_\mu, \cdots, a_1, 1, a_{-1}, \cdots, a_{-\nu}) \\ &\geqq k_1 \frac{L(1, \cdots, 1, a_\mu, \cdots, a_{-\nu})}{L(1, \cdots, 1)}.\end{aligned}$$

and hence $\mu(\sigma'_\lambda)$ is bounded away from zero, for arbitrary λ. Let $c > 0$ be such a lower bound.

Given $\epsilon = k_1 k_2^{-1} 3^{-1} c \mu(E \cdot D)$, there exists an open set E_0 such that $E_0 \supset E \cdot D$ and $\mu(E_0 - E \cdot D) < \epsilon$. For λ sufficiently large, we have from Case IV,

$$\mu(E_0 \cdot \sigma'_\lambda) > 2k_1 3^{-1} k_2^{-1} \mu(E_0) \mu(\sigma'_\lambda) \geqq 2k_1 3^{-1} k^{-1} \mu(E \cdot D) \mu(\sigma'_\lambda).$$

But

$$\mu(E_0 \cdot \sigma'_\lambda) - \mu(E \cdot \sigma'_\lambda) \leqq \mu(E_0 - ED) < \epsilon,$$

and hence

$$\mu(E \cdot \sigma'_\lambda) \geqq 2k_1 3^{-1} k_2^{-1} \mu(E \cdot D) \mu(\sigma'_\lambda) - \epsilon.$$

From the choice of ϵ, it follows that

$$\mu(E \cdot \sigma'_\lambda) \geqq k_1 3^{-1} k_2^{-1} \mu(E \cdot D) \mu(\sigma'_\lambda).$$

The desired lemma holds with $k_5 = k_1 3^{-1} k_2^{-1}$.

7. Metrical transitivity.

THEOREM 7. 1. (*Metrical transitivity*). *If E is a measurable set of π which is invariant under the group Γ_2, either $\mu(E) = 0$ or $\mu[C_\pi(E)] = 0$.*

It can be assumed that $\mu(E) > 0$. From Lemma 2. 1, $\mu(E \cdot D) > 0$. Let Δ be a rectangle of the net in D. From Lemmas 6. 1–6. 4, if λ is chosen sufficiently large,

$$\begin{aligned}\mu(E \cdot \Delta) \geqq \mu(E \cdot \sigma'_\lambda) &> k_4 \mu(E \cdot \sigma'_\lambda) \geqq k_4 k_5 \mu(E \cdot D) \mu(\sigma'_\lambda) \\ &> k_4^2 k_5 \mu(E \cdot D) \mu(\sigma_\lambda) > k_3 k_4^2 k_5 \mu(E \cdot D) \mu(\Delta),\end{aligned}$$

or

$$\mu(E \cdot \Delta) > k \mu(\Delta),$$

where $k = k_3 k_4^2 k_5 \mu(E \cdot D) > 0$.

But this implies $\mu(E \cdot D) = \mu(D)$. For if this were not the case, there would be a point of D at which the metrical density of the set $E \cdot D$ would be zero. If P is such a point, a sufficiently small square, S, with P as center lies entirely in D and $\mu(E \cdot S) < k\mu(S)$. A sequence of non-overlapping rectangles of the net in D can be so chosen that $S = F + \sum_{i=1}^{\infty} R_i$, where F is a zero set. For each of these rectangles (7.2) holds and hence

$$\mu(E \cdot S) = \sum_{i=1}^{\infty} \mu(E \cdot R_i) > k \sum_{i=1}^{\infty} \mu(R_i) = k\mu(S).$$

This contradiction implies $\mu(E \cdot D) = \mu(D)$.

The set $C_\pi(E)$ is then a measurable invariant set such that $\mu(D \cdot C_\pi E) = 0$. From Lemma 2.1, $\mu(C_\pi E) = 0$, and Theorem 7.1 holds.

Theorem 7.1 implies the metrical transitivity of the group with respect to the real axis. For if this were not true there would exist a measurable non-zero set, E_1, of the real axis such that E_1 would be invariant under the group Γ and $C(E_1)$, with respect to the real axis, would not be a linear zero set. But then the set $E_1 \times E_1$, consisting of those points (ξ, η) of π such that both ξ and η belong to E_1, would be a measurable non-zero set of π, invariant under Γ_2, such that neither $\mu(E_1 \times E_1) = 0$ nor $C_\pi(E_1 \times E_1) = 0$. This contradicts Theorem 7.1, and hence the group Γ must be metrically transitive with respect to the real axis.

Conversely, the group[7] G generated by the transformations

$$T^{(j)}: \qquad \xi = x + a_j, \qquad\qquad (j = 1, 2, \cdots),$$

for which $\lim a_j = 0$, $n \to \infty$, $a_j \neq 0$, is metrically transitive with respect to the real axis, but the group G_2 is evidently not metrically transitive with respect to the plane.

BRYN MAWR COLLEGE,
BRYN MAWR, PA.

[7] Martin, *loc. cit.*, p. 611.

SOME INTRINSIC AND DERIVED VECTORS IN A KAWAGUCHI SPACE.

By J. L. Synge.

1. *Introduction.* Let there be a space of N dimensions with coördinates x^i, and let there be a function F of

$$(1.1) \qquad t, x^{(0)i}, x^{(1)i}, \cdots, x^{(m)i},$$

where t is a parameter, and

$$(1.2) \qquad x^{(q)i} = d^q x^i/dt^q, \qquad x^{(0)i} = x^i.$$

We define F to be an invariant in the sense of tensor calculus. This does not imply the invariability of the functional form of F but simply that, when we employ a different coördinate system $\bar{x}^i$, we are to associate with it a function $\bar{F}$, given by the transformation of the arguments of F, t being unchanged: that is,

$$(1.3) \qquad \bar{F}(t, \bar{x}^{(0)i}, \cdots, \bar{x}^{(m)i}) = F(t, x^{(0)i}, \cdots, x^{(m)i}).$$

The length of a curve $x^i = x^i(t)$ from $t = t_1$ to $t = t_2$ is defined to be the invariant

$$(1.4) \qquad s = \int_{t_1}^{t_2} F dt.$$

Following Craig,[1] we shall call such a space a *Kawaguchi space of order m*, although Kawaguchi did not include t among the arguments of F.[2] If $m = 1$ in (1.1), if t is absent, and if F is homogeneous of the first degree in $x^{(1)i}$, the Kawaguchi space reduces to a Finsler space. If, more particularly, F is the square root of a homogeneous quadratic expression in $x^{(1)i}$, the Finsler space reduces to a Riemannian space. It will be noticed that we have made no reference to transformation of the parameter t. In both the Finsler space and the Riemannian space, the arc s as given by (1.4) is independent of the particular parameter t employed. In the Kawaguchi space, as discussed in the present paper, no condition is imposed on F to insure invariance of s under transformation of the parameter t. With the exception

[1] H. V. Craig, "On a generalized tangent vector," *American Journal of Mathematics*, vol. 57 (1935), p. 457.

[2] A. Kawaguchi, "Die Differentialgeometrie in der verallgemeinerten Mannigfaltigkeit," *Rendiconti Circolo Matematico di Palermo*, vol. 56 (1932), pp. 245-276.

of (5.30), the results established will be true if this invariance exists, but they do not require it.[3]

Since the parallel displacement of a vector, together with the associated ideas of absolute derivative and covariant derivative, play a fundamental part in Riemannian geometry, it is natural in studying these more general types of geometry to attempt to define parallel propagation and the associated operations in a way which reduces to the familiar definition when the space reduces to Riemannian space. The operation of absolute differentiation of a contravariant vector defined along a curve has been defined in the Finsler space by Taylor and Synge[4] and in the Kawaguchi space of the second order by Craig,[3] under the restriction stated. As I understand the work of Kawaguchi,[2] he appears to be interested in the most general forms which the operations in question could have, rather than the explicit development of the operations in terms of the function F. It is with this last development that the present paper is concerned.

Before proceeding to the discussion of absolute differentiation and parallel propagation along a curve, it is natural to develop the purely intrinsic properties of the curve itself. The vector which undergoes parallel propagation (unless *e.g.* it is a tangent vector) is not to be regarded as intrinsic.

For the Kawaguchi space of order m I develop a set of $m+1$ intrinsic covariant vectors associated with a curve.[5] When the space is Riemannian, $m=1$, and there are just two of these vectors, corresponding to the tangent and first normal. In addition to these vectors, denoted by $\overset{p}{E}_i$, other intrinsic vectors are also developed. The mode of development is based on taking some invariant "generating function" H of the variables (1.1), or a larger set containing derivatives of higher orders with respect to the coördinates.

Passing on to the definition of the absolute derivative of a vector along a given curve, it appears that the most natural process is that which derives

---

[3] In developing the theory of the Kawaguchi space for $m=2$, H. V. Craig ("On parallel displacement in a non-Finsler space," *Transactions of the American Mathematical Society*, vol. 33 (1931), p. 129) subjects F to the condition that s shall be invariant under transformation of t. Craig's method has been extended to general values of m by H. Hombu, "On a non-Finsler metric space," *Tôhoku Mathematical Journal*, vol. 37 (1933), pp. 190-198.

[4] J. H. Taylor, "A generalization of Levi-Civita's parallelism and the Frenet formulas," *Transactions of the American Mathematical Society*, vol. 27 (1925), pp. 246-264; J. L. Synge, "A generalization of the Riemannian line-element," *ibid.*, pp. 61-67. These papers were written simultaneously and independently.

[5] Three of these vectors have been given by Craig: see ref. 1.

a covariant vector from a contravariant vector given along the curve. We are naturally most interested in formulae which involve only the *first* derivatives of the components of the given vector. We find that there are $m+1$ such derived vectors for a Kawaguchi space of order m. One of these derived vectors may be picked out as the natural generalisation of the ordinary absolute derivative (whose vanishing implies parallel propagation), but it is interesting to note that the case $m=1$ is rather peculiar as far as the general formula is concerned.

2. *The Eulerian vector E_i.* In connecting with the variational equation

$$\delta \int_{t_1}^{t_2} F dt = 0, \tag{2.1}$$

there are associated the well-known Eulerian equations. It is natural to expect that the expressions which are equated to zero in these equations are the components of a covariant vector. That is in fact the case, and it seems most natural to use a variational method to prove it. Craig [1] has established it by a direct method.

Let us take a singly infinite family of curves,

$$x^i = x^i(u, t), \tag{2.2}$$

where t is a variable parameter along each curve and u is constant along each curve. Then

$$\frac{d}{du} \int_{t_1}^{t_2} F dt = \int_{t_1}^{t_2} \frac{\partial F}{\partial u} dt, \tag{2.3}$$

and this is an invariant. Following Craig, we shall adopt the convenient notation

$$F_{(q)i} = \frac{\partial F}{\partial x^{(q)i}}, \qquad F^{\cdot\cdot(p)}_{(q)i} = \frac{d^p}{dt^p} \frac{\partial F}{\partial x^{(q)i}}. \tag{2.4}$$

In the present connection d/dt means partial differentiation with respect to t, u being held fixed. Then

$$\begin{aligned} \frac{\partial F}{\partial u} &= \sum_{q=0}^{m} F_{(q)i} \frac{\partial}{\partial u} x^{(q)i} \\ &= \sum_{q=1}^{m} F_{(q)i} \frac{d}{dt} \frac{\partial}{\partial u} x^{(q-1)i} + F_{(0)i} \frac{\partial x^i}{\partial u} \\ &= \frac{d}{dt} \sum_{q=1}^{m} F_{(q)i} \frac{\partial}{\partial u} x^{(q-1)i} + F_{(0)i} \frac{\partial x^i}{\partial u} - \sum_{q=1}^{m} F^{\cdot\cdot(1)}_{(q)i} \frac{\partial}{\partial u} x^{(q-1)i}. \end{aligned} \tag{2.5}$$

Proceeding by successive steps in this way, we get

$$\frac{\partial F}{\partial u} = \frac{d}{dt} \{ \sum_{p=1}^{m} (-1)^{p-1} \sum_{q=p}^{m} F^{\cdot\cdot(p-1)}_{(q)i} \frac{\partial}{\partial u} x^{(q-p)i} \} + E_i \frac{\partial x^i}{\partial u}, \tag{2.6}$$

where

$$E_i = \sum_{q=0}^{m} (-1)^q F^{\cdot\cdot(q)}_{(q)i} ; \tag{2.7}$$

this is the Eulerian expression. We have then

$$\frac{d}{du}\int_{t_1}^{t_2} F dt = [\sum_{p=1}^{m} (-1)^{p-1} \sum_{q=p}^{m} F^{\cdot\cdot(p-1)}_{q(i)} \frac{\partial}{\partial u} x^{(q-p)i}]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} E_i \frac{\partial x^i}{\partial u} dt. \tag{2.8}$$

Now let us suppose that the curves (2.2) are so chosen that the points for which $t = t_1$ and $t = t_2$ are common to them all, and further that the values of

$$x^{(1)i}, x^{(2)i}, \cdots, x^{(m-1)i}$$

are also common to them all at these points. These conditions imply that

$$\frac{\partial}{\partial u} x^{(q)i} = 0 \text{ for } t = t_1 \text{ and } t = t_2, \qquad (q = 0, 1, \cdots, m-1). \tag{2.9}$$

Then the first term on the right-hand side of (2.8) vanishes, and we have

$$\frac{d}{du}\int_{t_1}^{t_2} F dt = \int_{t_1}^{t_2} E_i \frac{\partial x^i}{\partial u} dt. \tag{2.10}$$

Since this is an invariant, we have, on changing to new coördinates $\bar{x}^i$,

$$\int_{t_1}^{t_2} \left(\bar{E}_i \frac{\partial \bar{x}^i}{\partial u} - E_j \frac{\partial x^j}{\partial u}\right) dt = 0, \tag{2.11}$$

or

$$\int_{t_1}^{t_2} \left(\bar{E}_i - E_j \frac{\partial x^j}{\partial \bar{x}^i}\right) \frac{\partial \bar{x}^i}{\partial u} dt = 0. \tag{2.12}$$

But $\partial \bar{x}^i/\partial u$ is arbitrary along any one of the curves $u = \text{const.}$, except at the end points, where its components vanish; hence, by the usual method of the calculus of variations,

$$\bar{E}_i = E_j \frac{\partial x^j}{\partial \bar{x}^i}. \tag{2.13}$$

Hence we have the following result:

THEOREM I. *The Eulerian expressions,*

$$E_i \equiv \sum_{q=0}^{m} (-1)^q F^{\cdot\cdot(q)}_{(q)i} , \tag{2.14}$$

are the components of a covariant vector in a Kawaguchi space of order m.

Although not intrinsic, we may mention an invariant, which appears as

a by-product of the preceding work. We have seen that the last term on the right of (2. 8) is an invariant. Hence, if the terminal variations are left free, instead of being restricted by (2. 9), it follows that the first term on the right of (2. 8) is an invariant also. Let us put

$$X^i = \partial x^i/\partial u, \tag{2.15}$$

these being components of an arbitrary contravariant vector, given along any curve $u = \text{const.}$ The invariant in question is

$$\sum_{p=1}^{m} (-1)^{p-1} \sum_{q=p}^{m} F_{(q)i}^{\cdot\cdot(p-1)} (X^i)^{(q-p)}. \tag{2.16}$$

Writing it in a different form, we may state the following result:

THEOREM II. *The expression*

$$\sum_{r=0}^{m-1} (X^i)^{(r)} \sum_{p=1}^{m-r} (-1)^{p-1} F_{(p+r)i}^{\cdot\cdot(p-1)} \tag{2.17}$$

is an invariant in a Kawaguchi space of order m, X^i being the components of an arbitrary contravariant vector given along the curve $x^i = x^i(t)$.

When $m = 1$, this invariant is

$$F_{(1)i} X^i. \tag{2.18}$$

When $m = 2$, it is

$$X^i \{F_{(1)i} - \frac{d}{dt} F_{(2)i}\} + \frac{dX^i}{dt} F_{(2)i}. \tag{2.19}$$

3. *The set of intrinsic vectors $\overset{p}{E}_i$.* Since the establishment of the vector character of the Eulerian vector E_i, given in (2. 14), involves nothing beyond the fact that F is an invariant function of the variables (1. 1), it is obvious that we may obtain a class of vectors by the formula (2. 14) on substituting for F any function $f(F)$ of it. In Riemannian geometry it is convenient to substitute F^2. However, if the parameter t is chosen so as to make F constant along the curve, which can be done by taking for parameter

$$s = \int_{t_1}^{t} F dt, \tag{3.1}$$

it is easily seen that these vectors only differ from E_i by a constant factor.

It is to be borne in mind in all the subsequent work that new vectors may be obtained from those given by writing $f(F)$ instead of F.

Now let $\phi(t)$ be any function of t, transforming as an invariant on transformation of coördinates. Then ϕF is an invariant function of the

variables (1. 1), and we may use it as a "generating function" instead of F. Substituting in (2. 14), we deduce that

(3. 2) $$\sum_{q=0}^{m} (-1)^q (\phi F_{(q)i})^{(q)}$$

are the components of a vector: we have used the fact that ϕ involves t only. This reduces to

(3. 3) $$\sum_{q=0}^{m} (-1)^q \sum_{p=0}^{q} \binom{q}{p} \phi^{(p)} F_{(q)i}^{\cdot\cdot(q-p)} \quad ,$$

or

(3. 4) $$\sum_{p=0}^{m} \phi^{(p)} \overset{p}{E}_i,$$

where

(3. 5) $$\overset{p}{E}_i = \sum_{q=p}^{m} (-1)^q \binom{q}{p} F_{(q)i}^{\cdot\cdot(q-p)} \; .$$

Here $\binom{q}{p}$ is the usual binomial symbol,

(3. 6) $$\binom{q}{p} = \frac{q(q-1)\cdots(q-p+1)}{p!}, \quad \binom{0}{0} = \binom{q}{0} = 1,$$
$$(p, q = 1, 2, \cdots; q \geq p).$$

Now for any assigned value of t, we may choose the values of $\phi, \phi^{(1)}, \cdots, \phi^{(m)}$ arbitrarily, and they are all invariants. Let us make them all zero except $\phi^{(r)}$, and let $\phi^{(r)} = 1$. Then the vector (3. 4) reduces to $\overset{r}{E}_i$, and we may state the following result:

THEOREM III. *The expressions*

(3. 7) $$\overset{p}{E}_i = \sum_{q=p}^{m} (-1)^q \binom{q}{p} F_{(q)i}^{\cdot\cdot(q-p)}, \qquad (p = 0, 1, \cdots, m),$$

are the components of a set of $m+1$ covariant vectors in a Kawaguchi space of order m.[6]

It may be of interest to write out a few of these vectors explicitly:

(3. 8) $$\begin{cases} \overset{0}{E}_i = E_i = \sum_{q=0}^{m} (-1)^q F_{(q)i}^{\cdot\cdot(q)}, \\ \overset{1}{E}_i = \sum_{q=1}^{m} (-1)^q q F_{(q)i}^{\cdot\cdot(q-1)}, \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ \overset{m-2}{E}_i = (-1)^m \{F_{(m-2)i} - (m-1) F_{(m-1)i}^{\cdot\cdot(1)} + \tfrac{1}{2} m(m-1) F_{(m)i}^{\cdot\cdot(2)}\}, \\ \overset{m-1}{E}_i = (-1)^{m-1} \{F_{(m-1)i} - m F_{(m)i}^{\cdot\cdot(1)}\}, \\ \overset{m}{E}_i = (-1)^m F_{(m)i}. \end{cases}$$

[6] These vectors were obtained by Craig (ref. 1) for $p = 0$, $p = 1$, and $p = m$.

The vector character of the last is easily established directly.

For $m = 1$, there are only two vectors

$$\overset{0}{E}_i = F_{(0)i} - \frac{d}{dt} F_{(1)i}, \qquad \overset{1}{E}_i = - F_{(1)i}. \tag{3.9}$$

For $m = 2$, there are only three vectors

$$\begin{cases} \overset{0}{E}_i = F_{(0)i} - \dfrac{d}{dt} F_{(1)i} + \dfrac{d^2}{dt^2} F_{(2)i}, \\ \overset{1}{E}_i = - F_{(1)i} + 2 \dfrac{d}{dt} F_{(2)i}, \\ \overset{2}{E}_i = F_{(2)i}. \end{cases} \tag{3.10}$$

4. *The intrinsic vectors* $\overset{p,r}{G}_i$. The quantities $x^{(1)i} = dx^i/dt$ are components of a contravariant vector, and hence any one of the expressions

$$x^{(1)j}\overset{p}{E}_j, \qquad (p = 0, 1, \cdots, m), \tag{4.1}$$

is an invariant. In deriving the vector character of $\overset{p}{E}_i$ as in (3.7), all that was required was the fact the F is an invariant function of the variables (1.1). Now (4.1) is an invariant, but it is a function of

$$t, x^{(0)i}, x^{(1)i}, \cdots, x^{(2m-p)i}, \tag{4.2}$$

since these are involved in $\overset{p}{E}_j$. Thus we may use (4.1) as a generating function instead of F in (3.7), provided that we extend the range of summation to include the variables (4.2). Hence we have the result:

THEOREM IV. *The expressions*

$$\overset{p,r}{G}_i = \sum_{q=r}^{2m-p} (-1)^q \binom{q}{r} (x^{(1)j}\overset{p}{E}_j)_{\cdot\cdot(q)i}^{\cdot\cdot(q-r)}, \tag{4.3}$$
$$(p = 0, 1, \cdots, m;\ r = 0, 1, \cdots, 2m - p),$$

where

$$\overset{p}{E}_j = \sum_{s=p}^{m} (-1)^s \binom{s}{p} F_{\cdot\cdot(s)j}^{\cdot\cdot(s-p)}, \tag{4.4}$$

are the components of a set of $\frac{1}{2}(m+1)(3m+2)$ *covariant vectors in a Kawaguchi space of order* m.

Perhaps the most interesting of these vectors is that for which

$$p = m, \qquad r = m - 1. \tag{4.5}$$

We have

$$(4.6)\quad \overset{m,m-1}{G_i} = (-1)^{m-1}\{(x^{(1)j}\overset{m}{E_j})_{(m-1)i} - m(x^{(1)j}\overset{m}{E_j})^{\cdot\cdot(1)}_{(m)i}\}$$
$$= \delta_1{}^m F^{\cdot\cdot(1)}_{(1)i} - \delta_2{}^m F_{(2)i} - x^{(1)j}F_{(m)j(m-1)i} + m\frac{d}{dt}(x^{(1)j}F_{(m)j(m)i});$$

if $m = 1$, this becomes

$$(4.7)\quad \overset{1,0}{G_i} = F^{\cdot\cdot(1)}_{(1)i} - x^{(1)j}F_{(1)j(0)i} + \frac{d}{dt}(x^{(1)j}F_{(1)j(1)i}).$$

When the space is a Finsler space, F is homogeneous of degree unity in $x^{(1)i}$: then we have

$$(4.8)\quad \overset{0,1}{G_i} = F^{\cdot\cdot(1)}_{(1)i} - F_{(0)i},$$

which is the Eulerian vector, to within a sign. We may therefore state the following result:

THEOREM V. *In a Kawaguchi space of order m there is associated with each point of a curve a covariant vector* $\overset{m,m-1}{G_i}$ *given by* (4.6). *When the space is a Finsler space, this vector is identical with the Eulerian vector, except for sign.*

5. *The absolute derivative of a contravariant vector along a curve.* Let X^i be a contravariant vector field, the components being functions of the coördinates only. Let us take as generating function any one of the expressions

$$(5.1)\quad X^j\overset{p}{E_j} \qquad (p = 0, 1, \cdots, m).$$

This is a function of the variables (4.2), and hence may be substituted for F in (3.7), provided that the range of summation is suitably changed, the p of (3.7) being changed to another letter. In fact, we get a vector constructed as in (4.3). In order to show that the vectors obtained in this way are "derived" from X^j, we shall adopt the notation

$$(5.2)\quad \overset{p,r}{D_{ij}}X^j = \sum_{q=r}^{2m-p}(-1)^q\binom{q}{r}(X^j\overset{p}{E_j})^{\cdot\cdot(q-r)}_{(q)i},$$
$$(p = 0, 1, \cdots, m;\ r = 0, 1, \cdots, 2m - p).$$

We may state this result:

THEOREM VI. *The formula* (5.2) *defines a set of* $\frac{1}{2}(m+1)(3m+2)$ *covariant vectors derived (along a given curve) from a contravariant vector field* X^i *in a Kawaguchi space of order m,* $\overset{p}{E_j}$ *being as given in* (4.4).

If $r=0$, we shall have, in the course of the calculation, to take the partial derivative of X^j with respect to x^i. Consequently the derived vector will involve the partial derivatives of the vector field, as well as its derivatives with respect to t along the curve. But if we take $r>0$, the formula will involve only the derivatives of X^j with respect to t, as in the formula for the absolute derivative in Riemannian space, which is

$$\frac{dX^i}{dt}+\left\{ \begin{matrix} i \\ jk \end{matrix} \right\} X^j \frac{dx^k}{dt}. \tag{5.3}$$

The formula (5.2) will involve the derivatives of X^j with respect to t up to and including the order $2m-p-r$. Let us confine our attention to those which involve derivatives of the first order only. We are then to take

$$2m-p-r=1, \qquad r=2m-p-1. \tag{5.4}$$

Let us then consider the derived vectors

$$\overset{p,2m-p-1}{D_{ij}} X^j=(-1)^{p+1}\{(X^j\overset{p}{E}_j)_{(2m-p-1)i}-(2m-p)(X^j\overset{p}{E}_j)^{\cdot\cdot(1)}_{(2m-p)i}\}, \tag{5.5}$$
$$(p=0,1,\cdots,m).$$

If

$$2m-p-1=0, \tag{5.6}$$

then $\partial X^j/\partial x^i$ will appear in the evaluation of this expression. Now (5.6) can be true only if $m=1$, and then for the value $p=1$. It is interesting to see what (5.5) gives in this particular case. We have

$$\begin{aligned} \overset{1,0}{D_{ij}}X^j &= (X^j\overset{1}{E}_j)_{(0)i}-(X^j\overset{1}{E}_j)^{\cdot\cdot(1)}_{(1)i} \\ &=\frac{\partial X^j}{\partial x^i}\overset{1}{E}_j+X^j(\overset{1}{E}_j)_{(0)i}-\frac{d}{dt}\{X^j(\overset{1}{E}_j)_{(1)i}\}, \end{aligned} \tag{5.7}$$

where (cf. (3.8)) $\overset{1}{E}_j=-F_{(1)j}$: thus we have

$$\begin{aligned} \overset{1,0}{D_{ij}}X^j &=-\frac{\partial X^j}{\partial x^i}F_{(1)j}-X^jF_{(1)j(0)i}+\frac{d}{dt}(X^jF_{(1)j(1)i}) \\ &=\frac{dX^j}{dt}F_{(1)j(1)i}+X^j(F_{(1)j(1)i(1)k}x^{(2)k} \\ &\qquad +F_{(1)j(1)i(0)k}x^{(1)k}-F_{(1)j(0)i})-\frac{\partial X^j}{\partial x^i}F_{(1)j}. \end{aligned} \tag{5.8}$$

To complete the case $m=1$, we must also put $p=0$: this gives

$$\overset{0,1}{D_{ij}}X^j=-(X^j\overset{0}{E}_j)_{(1)i}+2(X^j\overset{0}{E}_j)^{\cdot\cdot(1)}_{(2)i}. \tag{5.9}$$

Here we are to put (cf. (3.8))

$$\overset{0}{E}_j = F_{(0)j} - F^{\cdot\cdot(1)}_{(1)j} = F_{(0)j} - F_{(1)j(1)k}x^{(2)k} - F_{(1)j(0)k}x^{(1)k}. \tag{5.10}$$

Thus

$$\begin{aligned} \overset{0,1}{D}_{ij}X^j &= -X^j(\overset{0}{E}_j)_{(1)i} + 2\frac{d}{dt}\{X^j(\overset{0}{E}_j)_{(2)i}\} \\ &= -X^j(F_{(0)j(1)i} - F_{(1)j(1)k(1)i}x^{(2)k} \\ &\quad - F_{(1)j(0)i} - F_{(1)j(0)k(1)i}x^{(1)k}) - 2\frac{d}{dt}(X^jF_{(1)j(1)i}) \\ &= -2[F_{(1)j(1)i}\frac{dX^j}{dt} + X^j\{\tfrac{1}{2}(F_{(0)j(1)i} - F_{(1)j(0)i}) \\ &\quad + \tfrac{1}{2}F_{(1)j(1)i(0)k}x^{(1)k} + \tfrac{1}{2}F_{(1)j(1)i(1)k}x^{(2)k}\}]. \end{aligned} \tag{5.11}$$

We may state this result:

THEOREM VII. *In a Kawaguchi space*[7] *of order* 1 *the formulae* (5.8) *and* (5.11) *define two covariant vectors derived from a contravariant vector field* X^j*:* (5.8) *involves the partial derivatives of* X^j*, but* (5.11) *involves only the derivatives with respect to* t.

It is interesting to see what the derived vectors (5.8) and (5.11) degenerate to in the case of a Finsler space. We shall not, however, use these formulae as they stand, but the corresponding formula with F^2 substituted for F. We shall denote the corresponding vectors by

$$\overset{1,0}{\Delta}_{ij}X^j, \qquad \overset{0,1}{\Delta}_{ij}X^j. \tag{5.12}$$

We shall write

$$f_{ij} = \tfrac{1}{2}(F^2)_{(1)i(1)j}, \qquad f_{ijk} = \tfrac{1}{2}(F^2)_{(1)i(1)j(1)k}. \tag{5.13}$$

We know that these are covariant tensors. Using the fact that F^2 is homogeneous of degree two in $x^{(1)i}$, we obtain

$$\begin{aligned} \overset{1,0}{\Delta}_{ij}X^j &= 2\left\{f_{ij}\frac{dX^j}{dt} + f_{ijk}X^jx^{(2)k}\right. \\ &\quad \left. + \left(\frac{\partial f_{ij}}{\partial x^k} - \frac{\partial f_{jk}}{\partial x^i}\right)X^jx^{(1)k} - f_{jk}\frac{\partial X^j}{\partial x^i}x^{(1)k}\right\} \\ &= 2\left\{\left(f_{ij}\frac{\partial X^j}{\partial x^k} + X^j\frac{\partial f_{ij}}{\partial x^k}\right)\right. \\ &\quad \left. - \left(f_{jk}\frac{\partial X^j}{\partial x^i} + X^j\frac{\partial f_{jk}}{\partial x^i}\right)\right\}x^{(1)k} + 2f_{ijk}X^jx^{(2)k}. \end{aligned} \tag{5.14}$$

---

[7] In the sense of the present paper, a Kawaguchi space of order 1 is not necessarily a Finsler space: for the Finsler space, F is homogeneous of degree unity in $x^{(1)i}$.

In the still more particular case of Riemannian space, we have

$$f_{ij} = g_{ij}, \tag{5.15}$$

which are functions of the coördinates only. Then we have

$$\overset{1,0}{\Delta}_{ij} X^j = 2\left(\frac{\partial X_i}{\partial x^k} - \frac{\partial X_k}{\partial x^i}\right) x^{(1)k}, \tag{5.16}$$

whose vector character is well known.

For the other derived vector in a Finsler space, we have

$$\begin{aligned} -\tfrac{1}{2}\overset{0,1}{\Delta}_{ij} X^j = 2f_{ij}\frac{dX^j}{dt} & \\ + X^j \left\{ \left(\frac{\partial f_{ik}}{\partial x^j} - \frac{\partial f_{jk}}{\partial x^i}\right) x^{(1)k} + \frac{\partial f_{ij}}{\partial x^k} x^{(1)k} + f_{ijk} x^{(2)k} \right\}, & \end{aligned} \tag{5.17}$$

$$\begin{aligned} \overset{0,1}{\Delta}_{ij} X^j = -4 \Big\{ f_{ij} \frac{dX^j}{dt} & \\ + \tfrac{1}{2}\left(\frac{\partial f_{ij}}{\partial x^k} + \frac{\partial f_{ik}}{\partial x^j} - \frac{\partial f_{jk}}{\partial x^i}\right) X^j x^{(1)k} + \tfrac{1}{2} f_{ijk} X^j x^{(2)k} \Big\}. & \end{aligned} \tag{5.18}$$

Hence we may state this result:

THEOREM VIII. *When the Kawaguchi space of order* 1 *is a Finsler space the covariant vectors* (5.8) *and* (5.11) *derived from a contravariant vector field X^j degenerate* (*when F is replaced by F^2*) *to* (5.14) *and* (5.18).

Except for the numerical factor -4, formula (5.18) expresses the covariant components of the absolute derivative of a contravariant vector in a Finsler space, as defined by Taylor and Synge.[4] Since the vanishing of this absolute derivative is taken to define parallel propagation in a Finsler space, it seems appropriate to adopt the following definition of parallel propagation in a Kawaguchi space of order 1, which, it is to be remembered, differs from a Finsler space only in so far as F is not necessarily homogeneous of order one in $x^{(1)i}$.

DEFINITION. *A vector X^j is propagated parallelly along a curve in a Kawaguchi space of order* 1 *when it satisfies the equations*

$$\overset{0,1}{\Delta}_{ij} X^j = 0, \tag{5.19}$$

where

$$\begin{aligned} -\tfrac{1}{2}\overset{0,1}{\Delta}_{ij} X^j = (F^2)_{(1)i(1)j}\frac{dX^j}{dt} + \tfrac{1}{2}\{(F^2)_{(1)i(0)j} - (F^2)_{(1)j(0)i}\} X^j & \\ + \tfrac{1}{2}(F^2)_{(1)i(1)j(0)k} X^j x^{(1)k} + \tfrac{1}{2}(F^2)_{(1)i(1)j(1)k} X^j x^{(2)k}. & \end{aligned} \tag{5.20}$$

Let us now return to the general Kawaguchi space of order m and the derived vectors given in (5.5). Let us put

$$p = m - 1, \tag{5.21}$$

so that the formula reads

$$\overset{m-1,m}{D}_{ij} X^j = (-1)^m \{ (X^j \overset{m-1}{E}_j)_{(m)i} - (m+1)(X^j \overset{m-1}{E}_j)^{\cdot\cdot(1)}_{(m+1)i} \}. \tag{5.22}$$

We have already investigated this for $m = 1$ (cf. (5.9)) and we have seen that, if we put F^2 in place of F, it reduces to a familiar expression in a Finsler space. Let us now reduce the expression for any value of m. We have

$$\overset{m-1,m}{D}_{ij} X^j = (-1)^m [X^j (\overset{m-1}{E}_j)_{(m)i} - (m+1) \frac{d}{dt} \{X^j (\overset{m-1}{E}_j)_{(m+1)i}\}]. \tag{5.23}$$

Let us turn to (3.8); we have

$$\begin{aligned} \overset{m-1}{E}_j &= (-1)^{m-1} (F_{(m-1)j} - m F^{\cdot\cdot(1)}_{(m)j}) \\ &= (-1)^{m-1} (F_{(m-1)j} - m \sum_{q=0}^{m} F_{(m)j(q)k} x^{(q+1)k}), \end{aligned} \tag{5.24}$$

and so

$$\begin{aligned} (\overset{m-1}{E}_j)_{(m)i} &= (-1)^{m-1} (F_{(m-1)j(m)i} \\ &\qquad - m F_{(m)j(m-1)i} - m \sum_{q=0}^{m} F_{(m)j(q)k(m)i} x^{(q+1)k}), \\ (\overset{m-1}{E}_j)_{(m+1)i} &= -(-1)^{m-1} m F_{(m)j(m)i}. \end{aligned} \tag{5.25}$$

Thus (5.23) reads

$$\begin{aligned} \overset{m-1,m}{D}_{ij} X^j &= -m(m+1) F_{(m)i(m)j} \frac{dX^j}{dt} \\ &\quad + (m F_{(m-1)i(m)j} - F_{(m)i(m-1)j} - m^2 \sum_{q=0}^{m} F_{(m)i(m)j(q)k} x^{(q+1)k}) X^j, \end{aligned} \tag{5.26}$$

which agrees with (5.11) when we put $m = 1$.

We get a derived vector analogous to (5.26) if we substitute for F any function of F. The earlier work indicates that F^2 is the most suitable function to take. We may state the following result:

THEOREM IX. *In a Kawaguchi space of order m the formula*

$$\begin{aligned} \overset{m-1,m}{\Delta}_{ij} X^j &= -m(m+1)(F^2)_{(m)i(m)j} \frac{dX^j}{dt} \\ &\quad + \{ m (F^2)_{(m-1)i(m)j} - (F^2)_{(m)i(m-1)j} \\ &\qquad - m^2 \sum_{q=0}^{m} (F^2)_{(m)i(m)j(q)k} x^{(q+1)k} \} X^j \end{aligned} \tag{5.27}$$

defines a covariant vector along any curve $x^i = x^i(t)$ along which the contravariant vector X^j is given.

The following definitions may be set down:

DEFINITION. $\overset{m-1,m}{\Delta_{ij}} X^j$ *is the absolute (covariant) derivative of the vector* X^i.

DEFINITION. *A vector X^j is propagated parallelly along a curve in a Kawaguchi space of order m if its components satisfy the differential equations*

$$\overset{m-1,m}{\Delta_{ij}} X^j = 0. \tag{5.28}$$

The type of geometry with which the present paper deals is essentially concerned with processes of generalisation. Generalisations are by no means unique, and a method such as that developed in the present paper opens up an embarrassing variety of generalisations. Riemannian geometry is the well-established base with which (as a particular case) our generalisations are to be checked. The definitions adopted above do check with well-established results in Riemannian space, but we may ask whether, in the case of the Kawaguchi space of order m, we have been wise to use F^2 as the generating function in (5.27), instead of F^{m+1}. This would equally well give agreement with results in Riemannian or Finsler space, since when $m = 1$ we have $m + 1 = 2$.

It is easily seen that

$$f_{ij} = \tfrac{1}{2}(F^2)_{(m)i(m)j} \tag{5.29}$$

is a covariant tensor in a Kawaguchi space of order m; if the determinant of f_{ij} is not zero, we may introduce a conjugate contravariant tensor f^{ij}, and use it to convert covariant vectors into contravariant vectors. Thus although the formula (5.27) derives a covariant vector from a contravariant vector X^j, we can at once obtain a contravariant derived vector,

$$\overset{m-1,m}{\Delta^k{}_{.j}} X^j = f^{ki} \overset{m-1,m}{\Delta_{ij}} X^j. \tag{5.30}$$

It should be noted, however, that this cannot be done if s is invariant under transformation of t, for in that case the determinant of f_{ij} vanishes.[8]

UNIVERSITY OF TORONTO.

[8] Cf. H. V. Craig, *Bulletin of the American Mathematical Society*, vol. 36 (1930), p. 560.

AN ANALYTIC CHARACTERIZATION OF SURFACES OF FINITE LEBESGUE AREA.[1] PART I.

By CHARLES B. MORREY, JR.[2]

Since Schwarz[3] showed that the ordinary definition of the length of a curve could not be generalized directly to give a definition of the area of a surface, many definitions of the area of a surface have been proposed. In this paper, we shall use that proposed by Lebesgue in his thesis.[4] Although this definition was almost forgotten for over twenty years due to the lack of methods for handling it and also perhaps for esthetic reasons, its usefulness in connection with the solutions of the Problem of Plateau (particularly those of Radó[5] and McShane[6]) demonstrates its value as a tool in Analysis and Geometry.

This definition presupposes a definition of limit elements in the field of surfaces. For surfaces $z = f(x, y)$, it is clear that we should say that a sequence of surfaces S_n had the surface S as its limit if and only if the corresponding functions $f_n(x, y)$ converged uniformly to the limiting $f(x, y)$. An ideal extension of this definition to general surfaces is furnished by Fréchet's definition of the distance between two surfaces.[7] It is a curious fact that, although earlier workers in the area of surfaces (such as Lebesgue and Geöcze) clearly had some such definition of convergent sequences in mind, it was not precisely formulated until so recently and has accordingly been used only in the work of Radó, McShane, Douglas, and the author.

The problem considered in this paper is that of determining an analytic

---

[1] Part I was presented to the American Mathematical Society on December 27, 1932, under the title "An analytic criterion that a surface possess finite Lebesgue area." Part II was presented on April 14, 1933, under its present title.

[2] National Research Fellow (1931-33).

[3] H. A. Schwarz, *Gesammelte Anhandlungen*, vol. 1, p. 309.

[4] H. Lebesgue, "Intégral, longueur, aire" (Dissertation), *Annali di Matematica*, ser. 3, vol. 7 (1902), pp. 231-359.

[5] T. Radó, "On the problem of least area and the problem of Plateau," *Mathematische Zeitschrift*, vol. 32 (1930), pp. 763-796.

[6] E. J. McShane, "Parametrizations of saddle surfaces with application to the problem of Plateau," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 716-733.

[7] M. Fréchet, "Sur la distance de deux surfaces," *Annales de la Société Polonaise de Mathematiques*, vol. 3 (1924), pp. 4-19.

characterization of surfaces of finite area more or less analogous to that of rectifiable curves. Accordingly, we shall list mainly researches on this problem and refer the reader to Radó's *Ergebnisse* tract "On the Problem of Plateau" [8] for the most important literature on the general theory of the area. Surfaces $z = f(x, y)$ of finite Lebesgue area were first characterized by Geöcze [9] and later independently by Tonelli.[10] Tonelli also characterized functions $f(x, y)$ (calling them absolutely continuous) for which the area of the surface $z = f(x, y)$ is given by the classical integral formula, and these functions have been invaluable in subsequent work on area. McShane [11] and the author [12] independently defined a class of representations (called class L) of surfaces for which $L(S)$ is finite and given by the classical integral formula and McShane [13] characterized "saddle" surfaces of finite area bounded by Jordan curves by showing that each such surface possesses a representation of class L (in fact "generalized conformal").

The present paper gives an analytic characterization of the most general surface of finite Lebesgue area. It is first shown (in part I) that every non-degenerate (see § 1) surface of finite Lebesgue area possesses a generalized conformal (see § 2) representation. To characterize arbitrary surfaces, it is found helpful to allow parametric representations of surfaces on certain sets in 3-space called *hemicactoids*, a theory of such representations having been fully developed in the author's recent paper "The topology of (path) surfaces" [14] which will hereafter be referred to as T. It is then shown in § 3 (part II) that a necessary and sufficient condition for a surface S to possess finite Lebesgue area is that there exists a hemicactoid $\bar{H}$ on which S may be represented, the representation being generalized conformal on each non-degenerate cyclic element (see § 3).

Throughout this paper we shall use the following vector notation: the

---

[8] T. Radó, "On the Problem of Plateau," *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (Springer), vol. 2 (1933).

[9] Z. de Geöcze, "Die notwendigen und hinreichenden Bedingungen für einer endlichen Flächeninhalt eines Flächenstückes," *Mathematicai es Physikai Lapok*, vol. 25 (1916), pp. 61-81.

[10] L. Tonelli, "Sulla quadratura delle superficie," *Atti della Reale Accademia dei Lincei*, ser. 6, vol. 3 (1926), pp. 357-362, 445-450, 633-638, 714-719.

[11] E. J. McShane, "Integrals over surfaces in parametric form," *Annals of Mathematics*, vol. 34 (1933), pp. 815-838.

[12] C. B. Morrey, Jr., "A class of representations of manifolds (Part I)," *American Journal of Mathematics*, vol. 55 (1933), pp. 683-707 (hereafter cited as R).

[13] *Loc. cit.* (first reference).

[14] C. B. Morrey, Jr., "The topology of (path) surfaces," *American Journal of Mathematics*, vol. 57, no. 1 (January, 1935), pp. 17-50.

letters x and X shall stand for the coördinates $(x^1, \cdots, x^N)$ and $(X^1, \cdots, X^N)$ of a point in the x-space in which the given surface lies, the letters u and U for (u, v) and (U, V) respectively, the sum and difference of pairs of these letters, i. e., $x_1 \pm x_2$ or $u_1 \pm u_2$, will denote the vector sum and differences in the respective spaces, x_α will stand for the vector $(\partial x^1/\partial\alpha, \cdots, \partial x^N/\partial\alpha)$, α being a parameter, $x(u)$ and $X(U)$ will be vector functions, and if ϕ is a vector in any space, $|\phi|$ shall denote its length. We shall sometimes write $x(P)$ to mean $x(u)$ where u is the coördinate vector of the point P. Given a point set E, $\bar{E}$ shall denote its closure and E^* the set of its frontier points. The letters r and R shall always denote Jordan regions (i. e., regions bounded by a single Jordan curve). All vector functions occurring in a transformation or a representation of a surface will be assumed to be continuous.

1. *Non-degenerate vector functions.* In this section, we shall merely demonstrate a few simple properties of such vector functions which, however, are invaluable in the developments of the next section.

Definition 1. Let $x(u)$ be a (continuous vector) function defined on $\bar{r}$. We define the *oscillation* of $x(u)$ over the set, E, as the least upper bound of $|x(u) - x(u')|$, for all u, u' in E. (T, def. 1, § 3.)

Definition 2. Let $x(u)$ be defined on $\bar{r}$ and suppose C is a continuum, in $\bar{r}$, of diameter $\geqq \rho > 0$. We define $\eta_1(\rho, x; C)$ as the oscillation of $x(u)$ over C and $\eta_1(\rho, x)$ the greatest lower bound of $\eta_1(\rho, x; C)$ for all such C.

Definition 3. We shall say that a continuum, C, is *the upper limit of a sequence*, $\{C_n\}$, of continua if (T, def. 1, § 2)

(i) all the limit points of a sequence, $\{P_n\}$, of points, $P_n \epsilon C_n$, lie on C;
(ii) if P is any point of C, there is a sequence, $\{P_k\}$, of points, $P_k \epsilon C_{n_k}$, which converges to P, $\{n_k\}$ being a subsequence of the integers.

If C is also the upper limit of every subsequence of $\{C_n\}$, then we say that C is the *limit* of the sequence $\{C_n\}$ and that $\{C_n\}$ *converges* to C.

The following lemma is well known:[15]

LEMMA 1. *If* $\{C_n\}$ *is a sequence of continua in a closed bounded region* $\bar{R}$, *then a subsequence of* $\{C_n\}$ *possesses a unique limit continuum*, C. *Thus the sets* (1) *of all continua of* $\bar{R}$, *and* (2) *of all continua in* $\bar{R}$ *of diameter* $\geqq \rho$, *are compact.*

[15] See, for instance, R. L. Moore, "Foundations of point set theory," *American Mathematical Society Colloquium Publications*, vol. 13, pages 28, 29.

THEOREM 1. *Suppose $x(u)$ is defined and continuous on $\bar{r}$. Then $\eta_1(\rho, x; C)$ is lower semicontinuous in C, and thus takes on its minimum, $\eta_1(\rho, x)$, on some continuum $\bar{C}$ of diameter $\geqq \rho$.*

Proof. Let $\{C_n\}$ be a sequence of continua, of diameter $\geqq \rho$, with limit continuum C. Let P_1 and P_2 be points of C such that

$$\eta_1(\rho, x; C) = | x(P_1) - x(P_2) |.$$

We may select a subsequence $\{n_k\}$ of the positive integers such that $P_i^{n_k} \to P_i$ $(P_i^{n_k} \epsilon C_{n_k})$ $(i = 1, 2)$, and $\eta_1(\rho, x; C_{n_k}) \to \varliminf_{n\to\infty} \eta_1(\rho, x; C_n)$. Then clearly

$$\begin{aligned}\eta_1(\rho, x; C) = | x(P_1) - x(P_2) | &= \lim_{k\to\infty} | x(P_1^{n_k}) - x(P_2^{n_k}) | \\ &\leqq \varliminf_{k\to\infty} \eta_1(\rho, x; C^{n_k}),\end{aligned}$$

which proves the theorem.

THEOREM 2. *If the (continuous vector) functions $x_n(u)$, defined on $\bar{r}$, approach $x(u)$ uniformly,*

$$\eta_1(\rho, x) \leqq \varliminf_{n\to\infty} \eta_1(\rho, x_n).$$

Proof. For each n, there exists a continuum C_n, of diameter $\geqq \rho$, such that $\eta_1(\rho, x_n) = \eta_1(\rho, x_n; C_n)$. We may select a subsequence, $\{n_k\}$, of integers such that $C_{n_k} \to C$ and $\eta_1(\rho, x_{n_k}) \to \varliminf_{n\to\infty} \eta_1(\rho, x_n)$. Now let P_1 and P_2 be points on C for which $| x(P_1) - x(P_2) |$ is a maximum, and let $\{n_l\}$ be a subsequence of the integers $\{n_k\}$ so that we can find points, P_{i,n_l} on C_{n_l}, so that $P_{i,n_l} \to P_i$, $(i = 1, 2)$. Then clearly

$$\begin{aligned}\eta_1(\rho, x) \leqq | x(P_1) - x(P_2) | &= \lim_{l\to\infty} | x_{n_l}(P_{1,n_l}) - x_{n_l}(P_{2,n_l}) | \\ &\leqq \lim_{l\to\infty} \eta_1(\rho, x_{n_l}) = \varliminf_{n\to\infty} \eta_1(\rho, x_n),\end{aligned}$$

which proves the theorem.

Definition 4. A vector function is said to be *non-degenerate* on a continuum, $\bar{C}$, if it is not constant over any continuum of $\bar{C}$ containing more than one point (cf. T, def. 5, § 4).

The following two theorems follow immediately from the definitions.

THEOREM 3. *If $x(u)$ is non-degenerate on $\bar{r}$, $\eta_1(\rho, x) > 0$ if $\rho > 0$.*

THEOREM 4. *If $x(u)$ is non-degenerate on $\bar{r}$, $u = u(U)$ is a $1 - 1$*

continuous transformation of $\bar{r}$ *into* $\bar{R}$, *and we define* $X(U) = x[u(U)]$, *then* $X(U)$ *is non-degenerate on* $\bar{R}$.

The following theorem simplifies the argument in § 2:

THEOREM 5. *If* $\{x_n(u)\}$ *is a sequence of non-degenerate vector functions approaching the non-degenerate vector function* $x(u)$ *uniformly, we can find a function* $\eta_1(\rho)$, *positive for* $\rho > 0$, *such that*

$$\eta_1(\rho) \leqq \eta_1(\rho, x); \qquad \eta_1(\rho) \leqq \eta_1(\rho, x_n), \qquad (n = 1, 2, \cdots).$$

Proof. We may define $\eta_1(\rho)$ as the greatest lower bound of $\eta_1(\rho, x)$ and the numbers $\eta_1(\rho, x_n)$. If this is zero for some $\rho > 0$, we may extract a subsequence, $\{n_k\}$, of the positive integers so that $\eta_1(\rho, x_{n_k}) \to 0$ which contradicts Theorems 2 and 3.

2. *The existence of a generalized conformal representation of an arbitrary non-degenerate surface of finite Lebesgue area.* In this section, we prove a selection theorem for a sequence of representations of non-degenerate surfaces which converge to a non-degenerate surface. This theorem together with its proof is the exact analog for the vector functions representing these surfaces of Lebesgue's selection theorem for a sequence of monotone functions with uniformly bounded Dirichlet integrals which converges uniformly on the boundary of a region. By means of this theorem, the main result of the paper is established. The method of proof used extends to representations on an n-sphere of n-dimensional manifolds which are of class L with

$$\int_{R_n} \cdots \int [g_{11} + \cdots + g_{nn}]^{n/2} du^1 \cdots du^n < M$$

independent of n, the g_{ii} being among the coefficients g_{ij} of the fundamental (positive definite) form

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} du^i du^j.$$

Definition 1. A function, $f(x, y)$, will be said to be *absolutely continuous in the sense of Tonelli*[16] (A. C. T.) in a region $\bar{r}$ if it is A. C. T. on every rectangle interior to r with f_x and f_y summable over r; $f(x, y)$ is A. C. T.[17] for $a \leq x \leq b$, $c \leq y \leq d$ if it is continuous there and

[16] L. Tonelli, "Sulla quadratura delle superficie," *Atti della Reale Accademia Nazionale dei Lincei*, ser. 6, vol. 3 (1926), pp. 633-638.

[17] In a paper, "Complements of potential theory II," *American Journal of Mathematics*, vol. 55 (1933), pp. 42-46, G. C. Evans has shown that this concept is identical with that of a continuous potential function of its generalized derivatives (see ref. 19).

(i) for almost every X, $a \leq X \leq b$, $f(X, y)$ is absolutely continuous in y, and for almost every Y, $c \leq Y \leq d$, $f(x, Y)$ is absolutely continuous in x;

(ii) $\int_a^b V_c^{(y)d}[f(X, y)]dX$ and $\int_c^d V_a^{(x)b}[f(x, Y)]dY$ both exist, where $V_c^{(y)d}[f(X, y)]$, for instance denotes the variation of $f(X, y)$ on (c, d) considered as a function of y alone.

It is known that f_x and f_y exist almost everywhere and are summable.

The following definitions and lemmas may be found in the literature and are included here merely for the sake of completeness.

Lemma 1.[18] *If $\{f_n(x, y)\}$ is a sequence of functions, A. C. T. on $\bar{r}$, which approach the continuous function, $f(x, y)$, uniformly and there exist constants M, $p > 1$, and $q > 1$, independent of n such that*

$$\iint_r [|\partial f_n/\partial x|^p + |\partial f_n/\partial y|^q]dx\,dy < M, \qquad (n = 1, 2, \cdots),$$

then $f(x, y)$ is A. C. T. on $\bar{r}$, $|f_x|^p$ and $|f_y|^q$ are summable on r, and

$$\iint_r |f_x|^p\,dx\,dy \leqq \varliminf_{n\to\infty} \iint_r |\partial f_n/\partial x|^p\,dx\,dy;$$

$$\iint_r |f_y|^q\,dx\,dy \leqq \varliminf_{n\to\infty} \iint_r |\partial f_n/\partial y|^q\,dx\,dy.$$

Lemma 2.[19] *Let $f(x, y)$ be A. C. T. in $\bar{r}_1$, and let $x = x(s, t)$, $y = y(s, t)$ be a $1 - 1$ transformation of $\bar{r}_1$ into $\bar{R}_1$ where $x(s, t)$ and $y(s, t)$ are continuous together with their first partial derivatives and $|x_s y_t - x_t y_s| \geq \lambda > 0$. Then if $\phi(s, t) = f[x(s, t), y(s, t)]$, we have that $\phi(s, t)$ is A. C. T. in $\bar{R}_1$ and*

$$\phi_s = f_x x_s + f_y y_s, \qquad \phi_t = f_x x_t + f_y y_t \tag{2.1}$$

almost everywhere.

Lemma 3. *Suppose* (i) *$f(x, y)$ is A. C. T. in $\bar{r}$ with f_x^2 and f_y^2 summable over r,* (ii) *$x = x(s, t)$, $y = y(s, t)$ is a $1 - 1$ conformal transformation (merely continuous on r^*) of $\bar{r}$ into $\bar{R}$,* (iii) *$\phi(s, t) = f[x(s, t), y(s, t)]$. Then* (a) *$\phi(s, t)$ is A. C. T. on $\bar{R}$,* (b) *its partial derivatives are given almost everywhere by the formulas* (2.1), (c) *ϕ_s^2 and ϕ_t^2 are summable over R, and* (d) *we have*

---

[18] C. B. Morrey, Jr., *loc. cit.* (R).

[19] G. C. Evans, "Fundamental points of potential theory," *Rice Institute Pamphlets*, vol. 7, no. 4 (1920), pp. 274-285, particularly.

$$\iint_R (\phi_s^2 + \phi_t^2)\,ds\,dt = \iint_r (f_x^2 + f_y^2)\,dx\,dy.$$

Proof. Lemma 3 is an immediate consequence of the preceding lemma as is easily seen by considering the mapping of regions entirely interior to R on regions entirely interior to r by the given transformation.

Definition 2.[20,21] A representation, $x = x(u)$, $u \epsilon \bar{r}$, of a surface, S, is said to be of class L if

(i) the components, $x^i(u, v)$, are all A. C. T. on $\bar{r}$, $(i = 1, \cdots, N)$,

$$\text{(ii)} \quad \lim_{h \to 0} \iint_{r_\alpha} \left| \frac{\partial(x_h^i, x_h^j)}{\partial(u, v)} - \frac{\partial(x^i, x^j)}{\partial(u, v)} \right| du\,dv = 0, \quad \alpha_0 > \alpha > 0,$$

$$(i, j = 1, \cdots, N),$$

$$x_h^i(u, v) = (1/h^2) \int_u^{u+h} \int_v^{v+h} x^i(\xi, \eta)\,d\xi\,d\eta,$$

r_α being the set of points of r at a distance $\geqq \alpha$ from r^* (i. e., this is true for *all* these α).

LEMMA 4.[20,21] *A convenient subclass of representation of class L is determined by the following conditions:*

(i) $x^i(u, v)$ *is A. C. T.*, $(i = 1, \cdots, N)$,

(ii) $|x_u^i|^p$, $|x_v^j|^q$ *are summable over* r, p, $q \geqq 1$, $1/p + 1/q \leqq 1$, $(i = 1, \cdots, N)$.

We include the case where one of p and q is unity and the other infinite by interpreting (ii), in the case where $p = 1$, $q = \infty$, for instance, to mean

(ii′) $|x_v^i| < M$, $|x_u^i|$ summable in r, $(i = 1, \cdots, N)$.

Surfaces $z = f(x, y)$ with $f(x, y)$ A. C. T. are also seen to be of class L.

LEMMA 5.[22,23] *If the representation,* $x = x(u)$, *of the surface S, is of class L, L(S) is given by the usual integral formula.*

Definition 3.[22] The representation, $x = x(u)$, of the surface S is *generalized conformal* if it satisfies conditions (i) and (ii) of Lemma 4 with

[20] C. B. Morrey, Jr., *loc. cit.* (R).
[21] E. J. McShane, *loc. cit.* (2nd ref., footnote 11).
[22] C. B. Morrey, Jr., *loc. cit.* (R).
[23] E. J. McShane, *loc. cit.* (2nd ref.).

$p=q=2$ and $E=G$, $F=0$ almost everywhere, E, F, G being given by their usual formulas.

Definition 4. We say that the points P_1 and P_2 of a surface S, $S: x=x(u)$, $u \in \bar{r}$, are *logically distinct* if they correspond to distinct values u_1 and u_2 in $\bar{r}$, such that $x(u)$ is not constant over any continuum containing them both. This property is clearly invariant under changes of parameter, u. If S[24] and $x(u)$ are non-degenerate, the above merely requires that $u_1 \neq u_2$.

LEMMA 6.[25] *Let* Π *be a non-degenerate polyhedron. It possesses a generalized conformal representation on the unit circle in which three given logically distinct points on the boundary of* Π *correspond to three given distinct points on the boundary of the unit circle. If* Π *is degenerate, the mapping is impossible.*

LEMMA 7.[26] *Let* S, $S: x=x(u)$, *and* S_n, $S_n: x=x_n(u)$, $(n=1,2,\cdots)$, *be continuous surfaces. Suppose* (i) *the given representations of the* S_n *are generalized conformal,* (ii) *the functions* $x_n(u)$ *converge uniformly to* $x(u)$, *and* (iii) $\lim L(S_n)=L(S)$. *Then the given representation of* S *is generalized conformal.*

THEOREM 1. *Given that* S, $S: x=x(u)$, $u \in \bar{r}$, *and* S_n, $S_n: x=x_n(u)$, $u \in \bar{r}$, $(n=1,2,\cdots)$, *are non-degenerate surfaces, that* $\lim_{n\to\infty} S_n=S$, *that* $x(u)$ *and* $x_n(u)$ *are all non-degenerate, and that* $x_n(u)$ *approaches* $x(u)$ *uniformly. Suppose* $x=X_n(U)$, $U \in \bar{R}$, *is a representation of* S_n *satisfying:* (i) *it is of class* L; (ii) *one of the induced* (T, § 4, Theorem 3, *and* Def. 8) *continuous monotone transformations,* $u=u_n(U)$, *of* $\bar{R}$ *into* $\bar{r}$ *carries three fixed (independently of* n) *distinct points,* A, B, *and* C, *of* R^* *into three fixed distinct points,* a, b, *and* c, *respectively, of* r^*; (iii) *there is a constant* M, *independent of* n, *such that*

$$\iint_R [\,|\partial X_n/\partial U|^2+|\partial X_n/\partial V|^2]\,dU\,dV < M, \qquad (n=1,2,\cdots).$$

Then the $X_n(U)$ *are equicontinuous on* $\bar{R}$.

If the $X_n(U)$ *are not normalized on the boundary (i. e., do not satisfy* (ii), *they are equicontinuous on any closed set interior to* R.

[24] A surface is said to be *non-degenerate* if it possesses a non-degenerate representation.

[25] See for instance, C. Caratheodory, "Conformal representation," *Cambridge Tracts in Mathematics and Mathematical Physics*, no. 28, § 161 and §§ 125-130 particularly.

[26] C. B. Morrey, Jr., "A class of representations of manifolds (Part II)," *American Journal of Mathematics*, vol. 56, no. 2 (1934), pp. 275-293.

In the normalized case, any limit function, $X(u)$, will satisfy all three conditions, and in the second case any limit function (defined over all of R) will satisfy (i) *and* (iii) *on every closed region interior to R.*

Proof. It is clear (Theorem 4, § 1) that we may take $\bar{r}$ to be the unit circle, and a, b, and c to be equally spaced. On account of Lemma 3, we may also take $\bar{R}$ to be the unit circle and A, B, and C to be equally spaced.[27] It is clearly sufficient to show that the functions $u_n(U)$ are equicontinuous.

We wish to observe at the outset that if P^*_1 and P^*_2 are points of R^* on a closed large (small) arc bounded by two fixed points and containing (not containing) the third, and we choose $\widehat{P^*_1P^*_2}$ as that arc bounded by P^*_1 and P^*_2 and lying in the above large (small) arc, then all the points of the arc $\widehat{P^*_1P^*_2}$ are carried into the corresponding arc $\widehat{p^*_1p^*_2}$, $p^*_i = T_n(P^*_i)$, $(i = 1, 2)$. This follows from the normalization of the T_n and the nature of the continua of R^* which are carried into points of r^* by a continuous monotone transformation.[28] Thus $| u_n(P^*_1) - u_n(P^*_2)|$ is equal to the oscillation of $u_n(U)$ on an arc $P^*_1P^*_2$ which contains at most one of the fixed points A, B, C, unless this oscillation is equal to 2 in which case the above expression is not less than $3^{1/2}$.

Let $C(P_0, \rho)$ denote the circle with any center at P_0 and radius ρ, $3^{1/2}/2 = d > \rho > 0$. Suppose that the oscillation of some $u_n(U)$ in $\bar{C}(P_0, \rho_0) \cdot \bar{R} \geqq \epsilon$, $2 \geqq \epsilon > 0$, $d > \rho_0 > 0$. Then, from T, § 3, Theorem 1,[28] it is clear that the oscillation of $u_n(U)$ on $[C(P_0, \rho) \cdot R]^* \geqq \epsilon$, $d > \rho \geqq \rho_0$. Define $C(\rho)$ to be the arc of $[C(P_0, \rho)]^*$ which lies in $\bar{R}$, and let $P^*_{1\rho}$ and $P^*_{2\rho}$ be its end points, if they exist, in which case we let $\widehat{P^*_{1\rho}P^*_{2\rho}}$ be the arc $[C(P_0, \rho) \cdot R]^* \cdot R^*$. Then it is clear that the oscillation of $u_n(U)$ over $C(\rho) \geqq \epsilon/2$, $d > \rho \geqq \rho_0$, for (i) if the oscillation over $P^*_{1\rho}P^*_{2\rho}$ (which may be null or a point) $\leqq \epsilon/2$, this is obviously the case, and (ii) if the oscillation over $\widehat{P^*_{1\rho}P^*_{2\rho}} > \epsilon/2$, then the oscillation over $C(\rho) \geqq | u_n(P^*_{1\rho}) - u_n(P^*_{2\rho})|$ which is $\geqq$ the smaller of the numbers $3^{1/2}$ and the oscillation of $u_n(U)$ over $\widehat{P^*_{1\rho}P^*_{2\rho}}$ (since $\widehat{P^*_{1\rho}P^*_{2\rho}}$ obviously cannot contain more than one of the fixed points), both of which exceed $\epsilon/2$, since $\epsilon \leqq 2$.

Now, by Theorem 5, § 1, we can find an $\eta_1(\epsilon)$, positive with ϵ, such that,

---

[27] The argument can be carried through if R is any (Jordan) region, however.

[28] Or it follows directly from the theorem (T. § 3, Theorem 2) that a monotone transformation is the uniform limit (in the sense that the vector functions approach their limit uniformly) of a sequence of $1 - 1$ continuous transformations, the statement being obvious for these.

for each $\epsilon > 0$, $0 < \eta_1(\epsilon) \leqq \eta_1(\epsilon, x)$, $0 < \eta_1(\epsilon) \leqq \eta_1(\epsilon, x_n)$, $(n = 1, 2, \cdots)$. Hence, if we define

$$\delta(\epsilon) = 2d \cdot k(\epsilon), \qquad k(\epsilon) = e^{-(2\pi M/[\eta_1(\epsilon/2)]^2)},$$

we see that $| u_n(U_1) - u_n(U_2) | < \epsilon$, when $| U_1 - U_2 | < \delta(\epsilon)$. For if this is not the case for some $u_n(U)$ and points U_1 and U_2, the oscillation of $u_n(U)$ in $\bar{C}(P_0, kd) \cdot \bar{R} \geqq \epsilon$, where $U_0 = (U_1 + U_2)/2$. Then for every ρ, $d > \rho \geqq kd$, the oscillation of $u_n(U)$ on $C(\rho) \geqq \epsilon/2$, and thus the oscillation of $X_n(U) = x_n[u_n(U)] \geqq \eta_1(\epsilon/2)$ on $C(\rho)$. Let us choose polar coördinates with pole at P_0 (notice Lemma 2), and let $\theta_{1\rho}$ and $\theta_{2\rho}$, $2\pi \geqq \theta_{2\rho} - \theta_{1\rho} > 0$ ($2\pi/3$ in fact), be the angular coördinates of $P^*_{1\rho}$ and $P^*_{2\rho}$ respectively (chosen so that $C(\rho)$ is the arc $\theta_{1\rho} \leq \theta \leq \theta_{2\rho}$) if they exist, otherwise let $\theta_{1\rho} = 0$, $\theta_{2\rho} = 2\pi$. Then using Schwartz's inequality

$$M > \iint_r \left[\left| \frac{\partial X_n}{\partial U} \right|^2 + \left| \frac{\partial X_n}{\partial V} \right|^2 \right] dU dV \geq \int_{kd}^{d} \frac{d\rho}{\rho} \cdot \int_{\theta_{1\rho}}^{\theta_{2\rho}} \left| \frac{\partial X_n}{\partial \theta} \right|^2 d\theta$$

$$\geqq \int_{kd}^{d} \frac{d\rho}{\rho} \cdot \frac{1}{\theta_{2\rho} - \theta_{1\rho}} \left[\int_{\theta_{1\rho}}^{\theta_{2\rho}} \left| \frac{\partial X_n}{\partial \theta} \right| d\theta \right]^2 \geq \frac{[\eta_1(\epsilon/2)]^2}{2\pi} \log \frac{1}{k} = M$$

which is impossible.

If we choose $r_0 < 1$ and $d = 1 - r_0$, the above argument demonstrates the equicontinuity of the $u_n(U)$ in the closed circle $U^2 + V^2 \leqq r_0^2$ independently of the normalization on $U^2 + V^2 = 1$. This demonstrates the second statement in the conclusion of the theorem. The third statement follows immediately from Lemma 1.

THEOREM 2. *A necessary and sufficient condition that a non-degenerate surface, S, be of finite Lebesgue area is that it possess a generalized conformal (normalized) representation on the closed unit circle.*

Proof. The sufficiency of the condition is immediate from Lemmas 4 and 5.

To prove the existence of such a map, let $\{\bar{\Pi}_n\}$ be a sequence of polyhedra approaching S, where $\lim_{n\to\infty} L(\bar{\Pi}_n) = L(S)$, and $x = x(u)$ be a non-degenerate representation of S on $\bar{r}$. It is clear that we may replace each $\bar{\Pi}_n$ by a non-degenerate polyhedron, Π_n, such that $| L(\bar{\Pi}_n) - L(\Pi_n) | < 1/n$ and $\| \Pi_n, \bar{\Pi}_n \| < 1/n$ by merely moving the vertices [29] of $\bar{\Pi}_n$ slightly. Then, let

[29] By definition (given for instance in C. B. Morrey, Jr., *loc. cit.* R) $\bar{\Pi}_n$ can be represented on $\bar{Q}$ (the unit square) by a function $\bar{x}_n(u)$ which is linear in triangles (a finite number of them). The vertices of $\bar{\Pi}_n$ are merely the points corresponding to the vertices of the triangles in $\bar{Q}$.

$x = x_n(u)$ be a sequence of non-degenerate representations of Π_n such that $x_n(u)$ approaches $x(u)$ uniformly (that this is possible follows from T, § 5, Theorem 2). Then, let a, b, and c be three distinct points of r^*, and A, B, and C three distinct points of R^*, where $\bar{R}$ is the closed unit circle. Let $x = X_n(U)$, $U \epsilon \bar{R}$, be a generalized conformal representation of Π_n on $\bar{R}$ so that an induced transformation, $u = u_n(U)$ of $\bar{R}$ into $\bar{r}$, carries A, B, and C into a, b, and c respectively. By Lemma 5, and the conformality,

$$L(\Pi_n) = (1/2) \int\int_r [| \partial X_n/\partial U |^2 + | \partial X_n/\partial V |^2] dUdV.$$

Thus, the hypotheses of Theorem 1 are fulfilled and thus we may extract a subsequence of the $X_n(U)$ which converges uniformly to a function $X(U)$. Clearly $x = X(U)$ is a representation of S [30] and, by Lemma 7, it is generalized conformal.

The following very important theorem due to McShane [31] and used by him in his very interesting solution of the problem of Plateau is a consequence of the above theorem and the theorem of T, § 5, Theorem 5.

THEOREM 3. *Let S be a Lebesgue monotone* (T, § 5, Def. 4) *surface of finite area bounded by a Jordan curve. Then S possesses a generalized conformal representation on the unit circle in which three given distinct points on the boundary of S correspond to three given distinct points on the circumference of the unit circle.*

THE UNIVERSITY OF CALIFORNIA.

---

[30] For let S_1 be this surface. Clearly $\lim_{n\to\infty} \| S, S_n \| = \lim \| S_1, S_n \| = 0$. But since the Fréchet distance satisfies the "triangle inequality," $\| S, S_1 \| \leqq \| S, S_n \| + \| S_1, S_n \|$, we see that $\| S, S_1 \| = 0$ and thus $S = S_1$.

[31] E. J. McShane, *loc. cit.* (1st ref.).